Special Report 66

U.S. ARMY RESEARCH INSTITUTE PROGRAM IN BASIC RESEARCH—FY 2007

Basic Research Unit
Paul A. Gade, Chief

May 2008

**U.S. Army Research Institute
for the Behavioral and Social Sciences**

**A Directorate of the Department of the Army
Deputy Chief of Staff, G1**

Authorized and approved for distribution:

**MICHELLE SAMS, Ph.D.
Director**

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) May 2008	2. REPORT TYPE Final	3. DATES COVERED (from. . . to) October 2002 to September 2007

4. TITLE AND SUBTITLE U.S. Army Research Institute Program in Basic Research - FY 2007	5a. CONTRACT OR GRANT NUMBER
	5b. PROGRAM ELEMENT NUMBER 611102A
6. AUTHOR(S) Basic Research Unit	5c. PROJECT NUMBER B74F
	5d. TASK NUMBER
	5e. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U. S. Army Research Institute for the Behavioral and Social Sciences 2511 Jefferson Davis Highway Arlington, VA 22202-3926	8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Institute for the Behavioral and Social Sciences 2511 Jefferson Davis Highway Arlington, VA 22202-3926	10. MONITOR ACRONYM ARI
	11. MONITOR REPORT NUMBER Special Report 66

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

Subject Matter POC: Dr. Paul A. Gade

14. ABSTRACT *(Maximum 200 words)*:

This document contains detailed summaries for each of the U.S. Army Research Institute's basic research contracts for the fiscal year 2007. These summaries are grouped according to four Basic Research Unit program objectives: Providing fundamental knowledge to improve training in complex environments; providing fundamental knowledge to improving leader and team performance; providing fundamental knowledge for identifying and measuring the attributes and skills that are critical to Soldier recruiting, assignment, performance, and retention in the transforming Army; and providing fundamental knowledge for organizational behavior and network science research. In addition to summarizing what was done or is being done, each summary also describes the contributions of that research effort to basic behavioral science and suggests how the findings might benefit the Army and other military services.

15. SUBJECT TERMS

Training, leadership, team performance, assessment, performance evaluation, adaptability, motivation, social networks, work behavior, network science

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited	144	Ellen Kinzer Technical Publication Specialist 703-602-8047

Standard Form 298

U.S. ARMY RESEARCH INSTITUTE PROGRAM IN BASIC RESEARCH - FY 2007

CONTENTS

CONTENTS (continued)

THE BASIC RESEARCH PROGRAM: IDENTIFYING
NEW AND PROMISING TECHNOLOGIES

The Basic Research Unit (BRU) research program focuses on providing the basic research underpinnings for the personnel, leader development, and training requirements of the future. This program is a critical link between the world of behavioral science and the military community. Searching out and advancing the state-of-the-art methods, theories, and findings in behavioral science; encouraging projects most likely to contribute generalizable scientific principles and new knowledge; and supporting those efforts that have potential military relevance and likelihood of leading to applied behavioral technology are BRU's key research goals.

Universities conduct most of the basic research in the program. BRU maintains close contact with other basic behavioral science research organizations as well as with ARI's applied researchers, other relevant agencies within the Army, and other military services. These contacts help to define issues that require fundamental research and facilitate the transition of basic research results to applied programs for eventual use by the operational Army.

In BRU's contract programs, a Broad Agency Announcement (BAA) is issued each year to solicit both concept papers and formal proposals relating to the announced program research areas. In a given year, the BAA highlights the research objectives of special interest and provides an open call for proposals.

On the following pages, the reader will find summaries of current and recently completed BRU contracts, which began between 2002 and 2007. There are four current BRU program objectives, each of which is discussed in detail below. These objectives are:

1. To provide fundamental knowledge to improve training in complex environments,
2. To provide fundamental knowledge to improve leader and team performance, and
3. To provide fundamental knowledge for identifying, assessing, and assigning quality personnel in the evolving Army, and
4. To provide fundamental knowledge for organizational behavior and network science research

Training in Complex Environments

Basic research in this area focuses on developing concepts and methods for training complex tasks and for sustaining complex task performance. This body of research also intends to assess the cognitive impact of technological requirements arising from digital, semi-automated, and robotic systems on training requirements. One of the efforts in this area seeks to find conditions that simultaneously optimize learning, memory, and transfer of training. This body of work also seeks to identify unique training principles and methods for improving interpersonal skills, team adaptability, and performance. One project evaluates an error-based approach in training complex interpersonal skills as opposed to the traditional behavioral modeling approach. The expected outcome under this objective is applied research testing the principles and methods produced from these efforts in Army training environments. The models

and theories produced should be useful in accounting for individual differences in training and facilitating practical, individualized, adaptive training methods. This body of research will lead to advances that will translate into improved training methods in a wide range of tasks.

Improving Leader and Team Performance

Commensurate with the requirements for rapidly developing adaptable, flexible leaders, the basic research program in improving leader performance is directed toward providing concepts and methods for accelerating leader development and understanding and developing leader adaptability and flexibility in a manner that can be tested in the applied environment. One of our efforts is focused on developing both adaptive leaders and teams through formal instruction, developmental work assignments, and self-development, while another project is identifying individual differences that are relevant in motivating leadership development. We are committed to discovering and testing the basic cognitive principles that underlie effective leader-team performance as well as understanding the dynamics of small group leadership in face-to-face and distributed team environments. One of these team-centered efforts evaluates how people's language, specifically the use of function words, relates to group dynamics. Results from this body of research will make an important contribution to understanding and improving organizational effectiveness through improving the ways we develop our leaders and teams.

Quality Personnel for the Evolving Army

Identifying and measuring the aptitudes and skills that are unique to the human performance requirements of military service is a major theme of this basic research effort. As part of this process, we seek to devise methods that assess sociological and psychological factors that could influence recruitment, retention, and Army performance. One effort is developing new multi-media assessments of emotional abilities and linking these measures to outcomes such as decision-making, teamwork, and leadership behaviors. Another project is developing a knowledge test to assess personality and temperament such that respondents cannot improve their scores through faking. Another effort is concerned with developing a valid, culture-fair test of intelligence that could be applied in selection and placement; another investigation examines the structure of citizenship and counterproductive work behavior and how it varies over time within individuals. Many of the basic research program's efforts fall under this objective, but all of these research efforts vary while still contributing to research on the selection, placement, and retention of personnel.

Organizational Behavior and Network Science

Research under this new objective is focused on using social networks, whether in simulations, games, or Army organizations to help understand and predict group behavioral processes. Modeling and understanding networks underlying human knowledge systems is also part of this objective. One effort under this objective intends to study what factors contribute to successful collaboration across distributed work groups. Another effort uses dynamic network analysis (DNA) to examine how Officers' organizational networks change over time. These and other efforts under this objective intend to study behavioral processes through the analysis of networks and changes within groups of individuals across a period of time.

This document provides a listing and brief synopsis of ongoing and recently completed research efforts. Project listings are organized into the three aforementioned research objectives. It is important to note, however, that basic research is but one of many programs for which BRU has responsibility.

Other programs in BRU include the following:

- The Small Business Innovative Research (SBIR) Program;
- The Small Business Technology Transfer (STTR) Program;
- The International Behavioral Science and Technology Watch;
- The graduate student apprenticeship program – the Consortium Research Fellows Program with the Consortium of Metropolitan Washington Universities; and
- Outreach efforts to Historically Black Colleges and Universities (HBCUs) and Minority Institutions (MIs), and research support in behavioral science for the U.S. Military Academy.

Additional information about reports from these research efforts is available upon request.

Paul A. Gade, Chief
Basic Research Unit

Tiffany M. Bludau, Consortium Fellow
Basic Research Unit

BRU RESEARCH OBJECTIVE #1: PROVIDE FUNDAMENTAL KNOWLEDGE TO IMPROVE TRAINING IN COMPLEX ENVIRONMENTS

Research under this objective develops concepts and methods for training complex tasks and sustaining complex task performance. The focus is on understanding the cognitive impact of Future Force technology on training requirements, the impact of voluminous multi-modal data on performance, and developing methods for improving cognitive skills.

Training Interpersonal Skills Using Error-Based Training

Contract #: W91WAW-07-C-0027   **Contract Dates:** 05/07/2007 to 05/06/2008
Institution: Personnel Decisions Research   **PI:** Michael J. Cullen
     Institutes (PDRI)

Problem(s)/Research Question(s) – Does an error-based approach to training complex interpersonal skills lead to superior learning and transfer outcomes compared to the traditional behavioral modeling approach?

Technical Barrier(s) – Interpersonal skills are difficult to train and assess. Assessment of the ability to perform interpersonal skills following training must be done over time and with multiple measures.

Significance/Impact for Basic Research – The project may contribute to a sound new theory and validated approach to training interpersonal skills.

Potential Transitions – Leadership skills training may profit from including error-based modules. In addition, performance in more technical military skill domains might be improved by using error-based training approaches.

Overview

For years, behavioral modeling has been the most popular method of training interpersonal skills. The popularity of behavioral modeling is due in large part to a series of early studies suggesting the efficacy of this approach in training such skills (e.g., Burke & Day, 1986; Latham & Saari, 1979; Meyer & Raich, 1983).

However, despite the popularity of behavioral modeling, and promising early studies, a large number of important questions remain unanswered about its effectiveness. One problem in assessing the effectiveness of behavioral modeling as an intervention is that many studies have used only paper and pencil situational judgment tests and reaction tests as outcome measures. Although these outcome measures are relevant, they are deficient because they do not allow for the crucial assessment of whether individuals can actually *perform* the interpersonal skills that have been trained. Behavioral measures, such as those based on role plays, are better outcome measures, because they allow for such an assessment. Unfortunately, when behavioral measures have been used as criteria, the results have not always supported the efficacy of behavioral modeling as a method of training interpersonal skills (Arthur, Bennett, Edens, & Bell, 2003).

One method for training interpersonal skills that holds promise is an "error-based" approach. Error-based learning is a relatively new approach to skill acquisition pioneered by cognitive psychologists, primarily in the educational domain (e.g. Brooks, 1990; Fosnot, 1996). The central premise of error-based training is that the learning of complex, cognitively-laden skills is best accomplished in an environment in which trainees actively engage in exploration, problem solving, hypothesis-testing, making mistakes, and learning how to recover from mistakes (Ivancic & Hesketh, 1995/1996). In contrast to a behavioral modeling approach, in

which the focus is on guiding learners in an errorless, step-by-step fashion through a pre-established set of training material, the focus in an error-based approach is on reducing training content, increasing participant involvement, and encouraging errors.

This increased cognitive load may lead to greater depth of processing of information. Greater depth of processing of information, in turn, is anticipated to lead to better encoding of information, and has been shown to lead to better recall (Craik & Lockhart, 1972). Another benefit of error-based training may be that it promotes more controlled, as opposed to automatic, processing of information (Ivancic & Hesketh, 1995/1996). Finally, error-based training may lead to enhanced learning outcomes because the learning environment in which such training takes place – typically a creative, less-structured, participative environment – may increase the motivation of learners.

Accordingly, this research tests the following hypotheses:

Hypothesis 1:   Trainees participating in an error-based training program for interpersonal skills, compared to trainees participating in a behavioral modeling program, will achieve higher knowledge acquisition and retention.

Hypothesis 2:   Trainees participating in an error-based training program for interpersonal skills, compared to trainees participating in a behavioral modeling program, will have better performance on behavioral demonstrations through role play.

Hypothesis 3:   Trainees participating in an error-based training program for interpersonal skills, compared to trainees participating in a behavioral modeling program, will have greater transfer of interpersonal skills to a novel interpersonal situation.

The research also tests the hypothesis, supported by recent research (Gully, Payne, Kiechel Koles & Whiteman, 2002), that learning will be optimized for trainees high in cognitive ability and openness to experience in the error-based condition, and for trainees lower on these dimensions in the behavioral modeling condition.

Research Approach

To test these hypotheses, 160 students will be recruited from the University of Minnesota-Twin Cities and randomly assigned to a behavioral modeling or error-based training program for teaching negotiation skills. Learning will be assessed using the three-prong framework suggested by Alliger, Tannenbaum, Bennet, Traver, and Shotland (1997). That framework includes assessment of knowledge acquisition, knowledge retention, and behavioral skill acquisition. In addition to these outcomes, transfer of skills to a different negotiation situation will be assessed. Immediately following training, participants will take the knowledge test and perform the behavioral skill role play. They will also complete Goldberg's 50-item International Personality Item Pool (1999) and sign a release form for their ACT/SAT scores.

One month following training, all participants will return to complete the transfer role play and the measure of knowledge retention.

Accomplishments

Data have not yet been collected for this effort.

Contributions to Basic Science

This project begins to address the question of how complex interpersonal skills, such as negotiation skills, ought to be trained to optimize relevant learning outcomes, including short- and long-term acquisition and retention of interpersonal knowledge, transfer of that knowledge to novel situations, and fluid performance of the interpersonal skills over long stretches of time. Our hypotheses suggest that the traditional method for training these skills, the behavioral modeling method, may not be the optimal method for achieving all of these outcomes. From a theoretical standpoint, the shortcoming of behavioral modeling may be that it relies too exclusively on the theory of observational learning as its guiding force, and does not sufficiently take into account advances in cognitive learning theory. These advances suggest that active learning and deep processing of information play a very important role in the acquisition, long-term retention, and performance of complex skills. If our main hypothesis is supported, this research will call into question continued reliance on observational learning as the appropriate theoretical paradigm for training interpersonal skills. It also will begin to lay the foundation for a validated new theory for training interpersonal skills that calls for much more active trainee participation at all stages of learning.

Potential Army/Military Applications

Because effective leaders must master a variety of interpersonal skills in order to motivate, mentor, inspire, support, and team build, this experiment may lead to important new advances in leader training that could ultimately enhance leader effectiveness. For instance, if our main hypothesis is supported, existing interpersonal skills training programs for officers (e.g., negotiation, communication, teamwork, and supervisory programs) could be modified to enhance the degree of active learning taking place. This would be expected to lead to increased leader acquisition and retention of these skills, and enhanced transfer and more fluid performance of these skills. Enhanced long-term performance of interpersonal skills should be important in non-combat and combat situations alike. To give one example, fluid communication and negotiation skills would be important in interactions with civilian and non-civilian populations overseas in times of combat. In addition to prompting a review and/or modification of current interpersonal skills training modules, the results of this research may warrant an investigation into whether training for many other more technical military skill domains could be enhanced by adopting a more active learning approach.

Future Plans

The training materials for this effort have now been created and finalized. In early 2008, we will gather the necessary data to test the hypotheses of this research effort.

References

Alliger, G. M., Tannenbaum, S. I., Bennett, W., Jr., Traver, H., & Shotland, A. (1997). A meta-analysis of relations among training criteria. *Personnel Psychology, 50,* 341-358.

Arthur, W., Bennett, W., Edens, P. S., & Bell, S. T. (2003). Effectiveness of training in organizations. A meta-analysis of design and evaluation features. *Journal of Applied Psychology, 88,* 234-245.

Brooks, J.G. (1990). Teachers and students: Constructivists forging new directions. *Educational Leadership, 47,* 68-71.

Burke, M. J., & Day, R. R. (1986). A cumulative study of the effectiveness of management training. *Journal of Applied Psychology, 71,* 232-245.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671-684.

Fosnot, C. T. (1996). Constructivism: A psychological theory of learning. In C.T. Fosnot (Ed.), *Constructivism: Theory, perspectives and practice.* New York: Teachers College Press.

Gully, S. M., Payne, S. C., Kiechel Koles, K. L., & Whiteman, J. K. (2002). The impact of error training and individual differences on training outcomes: An attribute-treatment interaction perspective. *Journal of Applied Psychology, 87*, 143-155.

Ivancic, K., & Hesketh, B. (1995/1996). Making the best of errors during training. *Training Research Journal, 1,* 103-125.

Latham, G. P., & Saari, L. M. (1979). Application of social-learning theory to training supervisors through behavioral modeling. *Journal of Applied Psychology, 64,* 239-246.

Meyer, H. H., & Raich, M. S. (1983). An objective evaluation of a behavioral modeling program. *Personnel Psychology, 36,* 755-762.

Factors Influencing the Design and Conduct
of Effective Technology-Delivered Instruction

Contract #: DASW01-04-K-0002
Institution: The University of Tulsa
               Illinois Institute of Technology

Contract Dates: 03/01/2004 to 09/30/2006
PI: Joanne Davis
Co-PI: Annette Towler, Kurt Kraiger
Liaisons: Scott Graham, IFRU;
Jim Belanich, ARI-PBS

Problem(s)/Research Question(s) – Seductive details refer to entertaining and interesting information inserted into training material that is actually tangential to the main theme of the topic. While previous research has shown that the use of seductive details and on-screen text hinders learning in a computer-based training environment, our research shows that these factors at times have no effect or a positive effect on learning. Our research centers on characteristics of learners and the learning environment that lead to either facilitative or inhibiting effects of seductive details.

Technical Barrier(s) - Previous research establishing the seductive detail effect has used a limited set of training materials, has focused exclusively on declarative knowledge, and has not adequately distinguished seductive details that are independent of or dependent of the primary training content. Little is known about the effects of prior knowledge of course content or trainees' learning styles. Finally, preliminary research suggests that seductive details are most likely to have an effect when learners approach their limits for cognitive load (how much information they can process at once), and these limits may vary from learner to learner.

Significance/Impact for Basic Research – The results will allow us to assess the extent to which training design characteristics impair both trainees' recollection of the facts and principles taught in training and their ability to perform the skills that were taught. In the long-term, this research will enable the military to design more effective computer-based training.

Potential Transitions –

- Basic and applied research on learning and instructional principles
- Instructional design for computer-based training/Web-based training
- Game-based training; determining whether details related to game fidelity detract from learning in game/simulation environments

Overview

One common strategy for enhancing any type of training is to make the content as interesting as possible. Intuitively, these strategies should increase intrinsic motivation in learners, resulting in greater appreciation for the material and greater persistence in learning environments. Obviously, interesting information that is central to the main topic is useful and can facilitate trainees' learning. However, trainers or instructional designers tend to include interesting information that is unrelated to the central theme to spice up mundane information.

This phenomenon is called the "seductive details effect" and refers to essentially "highly interesting and entertaining information that is only tangentially related to the topic but is irrelevant to the author's intended theme" (Harp & Mayer, 1998, p. 1). Instructional material can be seductive through inclusion of illustrations or text that is interesting and entertaining but tangential to the topic (e.g., adding a picture of a professional football team to a training module on teamwork).

Although one could argue that the inclusion of seductive details (SDs) increases trainees' emotional interest in the topic, there is evidence suggesting that the inclusion of this material disrupts trainees' concentration on the main themes of the material (Harp & Mayer, 1998; Mayer, Heiser, & Lonn, 2001). Thus, trainees learn less in courses where seductive details are included along with the presentation of material relevant to the topic of the course.

Previous seductive details research has focused exclusively on declarative knowledge (i.e., knowledge of the facts and principles taught in training), and most studies have used only a single set of instructional materials (e.g., Harp & Mayer, 1997, 1998; Mayer et al., 2001). In previous research described, we found that seductive details had no effect on recall but had a beneficial effect on transfer.

Research Approach – Experiment 1

Ninety students from two universities participated in a 10 minute computer-based training course. Participants ranged from 18-29 years of age with an average age of 21 years. The course demonstrated several functions in Excel that could be used to select a city to live in based on the city's average rainfall and temperature. Participants were randomly assigned to one of four experimental conditions: (1) training that included both seductive details and on-screen text, (2) training that included seductive details but not on-screen text, (3) training that included on-screen text but not seductive details, (4) training did not include seductive details or on-screen text. On-screen text summarized key points covered in training, while seductive details included interesting facts about the data used in training such as "Tree crickets are called the poor man's thermometer because temperature directly affects their rate of activity. Listen for a cricket and count the number of chirps it makes in fifteen seconds. Add 37. The sum will be the Fahrenheit temperature (almost exactly!)."

Accomplishments – Experiment 1

Use of Excel prior to training accounted for a significant 8% of the variance in procedural knowledge exam scores. However, seductive details and text summaries did not account for a significant portion of the variance in declarative knowledge ($R^2 = .003$, $p > .05$), procedural knowledge ($R^2 = .029$, $p > .05$), or training reactions ($R^2 = .035$, $p > .05$), after controlling for previous experience with Excel. Thus, including seductive details and/or text summaries did not affect training reactions or how much trainees learned, contrary to our expectations.

Seductive details and on-screen text interacted with trainees' learning styles. Including on-screen text greatly decreased performance for trainees with an auditory learning style and slightly decreased performance for trainees without an auditory learning style (see Figure 1).

This effect occurred for both declarative and procedural knowledge learning outcomes. Also, including seductive details greatly decreased performance on the procedural knowledge exam for trainees with a tactile learning style and slightly decreased performance for trainees without a tactile learning style (see Figure 2).
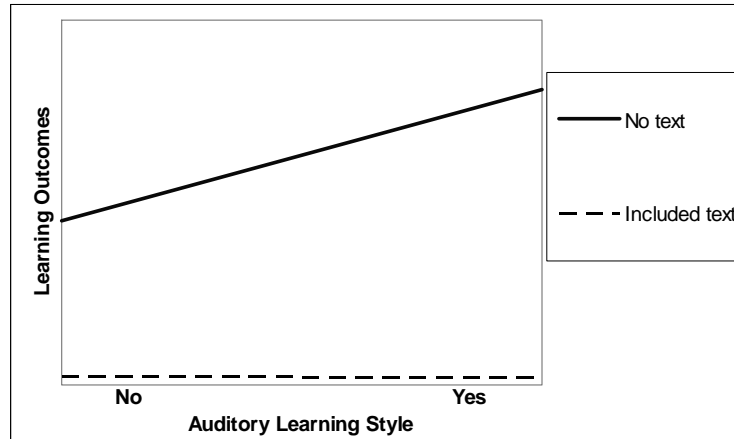
Figure 1. *Interaction between text summaries and auditory learning style when predicting declarative and procedural knowledge test scores.*
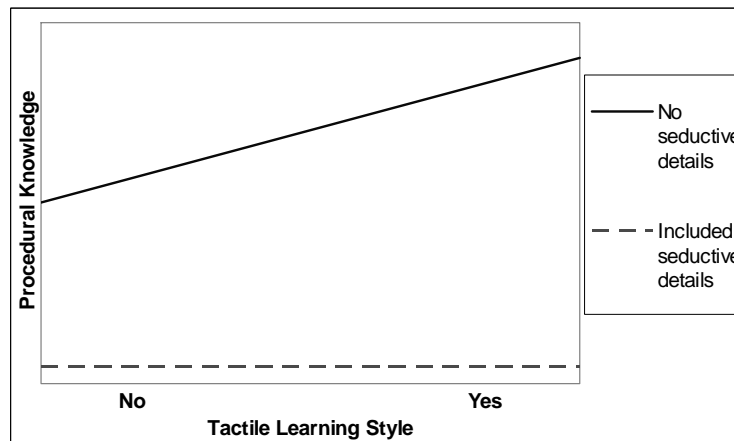
Figure 2. *Interaction between seductive details and tactile learning style when predicting procedural knowledge test scores.*

Research Approach – Experiment 2

The second experiment was a replication and extension of the first experiment. We attempted to replicate the findings from Experiment 1 in a different setting using another training program (Microsoft Mailmerge). We extended the second experiment through testing the effects of different types of SDs on recall and transfer performance. To better understand the seductive details effect, we tested whether SDs had a deleterious effect depending on whether they were independent or dependent on instructional content. For example, in the context of providing instruction on the use of Microsoft Mailmerge, an example of dependent seductive detail is "The

Windows NT software was developed by Microsoft-Israel," whereas an independent seductive detail is "Top graduate schools often accept up to only 10 applicants per year." As can be seen in the examples, dependent SDs are tangentially related to the instructional material, whereas independent SDs have no relevance to the instructional material.

Seventy-seven participants were recruited for this effort from a private Midwestern university. The independent variable consisted of three levels: no seductive details, independent seductive details, and dependent seductive details. As in Experiment 1, post-training measures were collected on a multiple-choice test of declarative knowledge and a performance-based measure of transfer after a time delay and distracter task.

Accomplishments – Experiment 2

As in Experiment 1, there was no main effect for seductive detail condition on declarative knowledge test scores. There was a main effect for seductive detail condition on transfer performance. Transfer performance was better in groups that were exposed to independent seductive details for the transfer task score ($M = 15.62$, $SD = 1.03$) compared to the control group ($M = 11.15$, $SD = 1.17$). There was no difference between the dependent seductive details ($M = 13.03$, $SD = 1.07$) and the other groups. Consistent with Experiment 1 (and contrary to the findings of many previous studies), the introduction of (independent) seductive details had a facilitative effect on transfer performance.

Research Approach – Experiment 3

Since we conducted two experiments with results that were very different than those obtained by Mayer and colleagues (e.g., Harp & Mayer, 1997, 1998; Mayer et al., 2001), we attempted to conduct a tight replication of one of Mayer's earlier studies. Sixty-two undergraduates at a Southwestern university participated in the investigation for extra-credit. Following the design in Harp and Mayer (1998), participants were assigned to one of four conditions: (1) a baseline condition, (2) baseline plus seductive text, (3) baseline plus seductive illustrations; and (4) baseline plus seductive illustrations and seductive text. In all conditions, participants read a scientific explanation of how lightning forms. Seductive details were in the form of interesting information about lightning but irrelevant to its formation (e.g., the number of Americans per year struck by lightning) and/or interesting pictures irrelevant to lightning formation (e.g., a high school football player whose uniform had been burned by lighting. All text and seductive details were identical to those used by Harp and Mayer. Participants completed the training material and answered four open-ended questions applying what they had learned. Participant responses were scored by two researchers and summed across questions. Note that these transfer questions are actually problem-solving or applied questions rather than performance-based measures as used in Experiments 1 and 2.

Accomplishments – Experiment 3

Two one-way ANOVA revealed significant main effects for seductive details on both the number of main points recalled and scores across the four "transfer" or problem-solving questions. Post hoc comparisons of recall measures indicate that participants recalled the most main points in the base (no seductive detail) condition, and the fewest main points in the base +

text + illustration condition. While post hoc comparisons revealed that problem-solving scores were lowest in the base + text + illustration condition, there were no differences between the other three conditions. In general, these results are similar to those obtained by Harp and Mayer (1998), although not as strong for the seductive detail text and seductive detail illustration conditions. Interestingly, Harp and Mayer concluded that "seductive details do their damage" by interfering with the organization of information at the point of learning. However, visual inspection of the free recall measures in the three seductive details conditions shows strong organization of information. Even though, upon presentation, the seductive details were interspersed with the primary information, at recall participants frequently presented the seductive details (organized) together at either the beginning or end of their recall statements.

Research Approach – Experiment 4

In a phone conversation with Richard Mayer, Dr. Mayer speculated that one difference between some of our experiments and his was that our experiments provided less cognitive load on participants, and thus seductive details were less disruptive. For example, when learning new software, there may be a natural organization to the information or learners can rely on prior knowledge of other software to organize the new information. Seductive details could be less disruptive under these conditions than if all information was new. Seductive details in general would be predicted to have the greatest effect when greater cognitive load was greatest. Dr. Mayer also speculated that there could be differences in the nature of our seductive details (from his). Over time, he had found that visually graphic or violent details tended to be the most disruptive.

To examine the effects of different levels of cognitive load and different types of seductive details, we conducted a fourth experiment. Sixty research participants from a private southwestern university were recruited for the experiment. They were assigned to one of four cells in a 2 x 2 design. Participants read about how a car engine works. Participants either read one page at a time or two pages at a time (with the latter condition requiring them to store more information at one time, thus imposing greater cognitive load). Also, participants either read and saw neutral seductive details or graphic ones (e.g., a car from a violent automobile accident, or female models posing at a new car show). We used a multi-item declarative knowledge measure to assess learning, and participants also rated their satisfaction with the training materials.

Accomplishments – Experiment 4

There was no effect for cognitive load (one vs. two pages) or seductive detail type (neutral vs. graphic) on the knowledge test scores. Participants reading two pages at a time rated the materials significantly more confusing than did participants reading one page at a time. In sum, there was little support for either manipulation on learning, although the requirement to read two pages at a time may have created more cognitive load for learners.

Research Approach – Experiment 5

Seventy-one participants were recruited for this effort from a private Midwestern university. The students ranged from freshman to undergraduate seniors. Twenty-seven

participants were females and 44 were males. Nearly all participants stated they were not familiar with Microsoft Mailmerge. Each participant was randomly assigned to one of four groups: (1) cognitive load + no SDs, (2) the control—no cognitive load + no SDs, (3) cognitive load + SDs, and (4) no cognitive load+ SDs.

The training provided instruction to participants on Microsoft Mailmerge—a function of the popular software package, Microsoft Word. The training session was an audio-visual file that participants listened to with headphones and viewed on a 17-inch computer monitor. The file contained animated videos with screen shots of Mailmerge and Excel programs. The overall topic of the training was to use Mailmerge to organize and send personalized letters in mass mailings. The training covered managing database information, filtering and sorting information, and inserting conditional statements. Cognitive load was introduced through use of a clicker. In the cognitive load condition, the participant clicked the clicker each time the word "document" was heard during the training.

Demographics were taken before the training. Motivation was measured through an 8-item survey with questions like "I am motivated to learn the skills emphasized in the training program." A 7-item reaction measure included questions like "I found the training coherent." A 13-question, multiple choice post test was administered to measure declarative knowledge. Transfer of training was measured through a task that included filtering, sorting, and merging of letters.

Accomplishments – Experiment 5

An ANCOVA was conducted to test for an interaction between cognitive load and seductive details with relation to transfer task scores. There was a significant interaction between cognitive load and seductive details ($p < .05$). An ANOVA was then conducted to further test for group differences in training outcomes. A post hoc test (Tukey HSD) revealed that participants who experienced cognitive load + SDs scored significantly ($p=.04$) lower on transfer than participants with no cognitive load + no SDs.

Contributions to Basic Science

The results from Experiments 1 and 2 suggest that seductive details can be effective for transfer. Prior research has focused primarily on the effects of seductive detail on recall, finding a negative effect. It is interesting to learn that seductive details can have a positive effect on task performance following a delay from training.

Results across all research efforts suggest that seductive details may have an effect on recall measures, but these effects do not occur when recognition (multiple-choice tests) are used. Researchers need to exercise care in choosing the appropriate dependent variable, and theories on seductive details should account for differential effects based on test type.

Visual inspection of participant responses in Experiment 3 suggests that if seductive details have negative effects on learning, they do not do so by preventing organization of new information.

Evidence from Experiment 5 (but not Experiment 4) suggests that increasing cognitive load may make seductive details more disruptive to learning.

Cognitive load theory suggests that learners can absorb a limited amount of material into working memory and if they are overloaded with information then they fail to form knowledge schemas that can be transmitted to long-term memory (Sweller, 1994). Cognitive load may be influenced by learner characteristics (e.g., cognitive ability or learning style), depth and complexity of content, or instructional design characteristics. There may be crucial "tipping points" for cognitive load to occur and thus for seductive details to do their damage. This point also may be dependent on individual differences, which we did not investigate in our research.

Potential Army/Military Applications

The knowledge acquired from this basic research emphasizes the importance of monitoring and controlling levels of cognitive load in computer-based training studies. Cognitive load thus emerges as a key variable for understanding relations among learner capabilities, instructional design features, and learning during training. Seductive details are but one mechanism for increasing cognitive overload. A logical next step is to study more broadly multiple instructional strategies that enhance learning and transfer by decreasing the learner's cognitive load. This research can be followed up by testing various design principles rooted in cognitive psychology in terms of their impact on cognitive load in various forms of computer-based training of work-related skills.

Understanding variables that affect learners' ability to separate understanding of seductive details from their understanding of core content may have other applications for the military as well. For example, in game-based training, seductive details take the form of graphics and special effects that enhance the appeal of the game but might be unrelated to the purpose of training. Seductive details and cognitive load theory may provide a platform for studying boundaries of game fidelity. Alternatively, identifying learner factors related to the capacity to "separate wheat from chaff" might have implications for designing effective training of intelligence analysts.

Final Summary

The purpose of the research was to examine how instructional characteristics (e.g., how information is presented) might affect trainee interest, trainee learning, and trainee transfer to task performance. Specifically, the research looked at the use of seductive details in instructional material; seductive details refer to interesting information that is irrelevant to the primary learning objectives (e.g., colorful "war stories" from the trainer's career).

Unlike past research by other researchers finding that the inclusion of seductive details has a negative effect on immediate recall, most of our research efforts found no negative effect for seductive details on learning: That is, there is as much learning when seductive details are present as when they are not. However, we also used a different type of learning measure (recognition-based, multiple choice questions) than have been used in prior research (recall-based, open-ended questions). A second major finding was that the inclusion of seductive details

had a positive impact on transfer performance—trainees performed better on a subsequent task following seductive details, presumably because the seductive details encouraged greater effort or deeper learning. A third major finding was that negative effects for seductive details on learning occurred only when trainees' cognitive load was high. In other words, when the learning is more difficult (either because of the content or the learning medium), seductive details are likely to be more distracting than when learning is straightforward.

These results have implications for the Army in terms of the design and delivery of effective training programs. In general, the more difficult the training, the more careful instructional designers or trainers should be about including ancillary material not directly related to the core learning objectives. For example, personal examples or cartoons embedded in viewgraphs might not matter on more straightforward content, but could distract learners on more difficult training content, or when the delivery of training is suboptimal (e.g., online material spread out over multiple small sessions). On the other hand, the more trainees will be expected to apply training content to field problem-solving scenarios, the more important it is for designers and trainers to think about ways of making learning purposeful.

The effects of cognitive load on learning were minimally explored in the current research effort. Future research should focus more extensively on not only the effects of cognitive load on the seductive details effect, but the extent to which other types of instructional strategies (e.g., providing learning objectives, using game-like practice scenarios) not only influence learners' cognitive load, but also their learning during training and their subsequent transfer performance.

More generally, the results of these experiments suggest that in all future training research conducted by the Army, multiple measures of learning and transfer should be collected. Interventions that affect one type of learning outcome might not affect others, and the link between learning in training and subsequent transfer of training is not always direct.

In terms of applied research, it is important that research be conducted to identify specific design features that influence the effectiveness of existing Army training programs. Most of the instructional design research (e.g., massed vs. distributed practice) was originally conducted by the military, but most of that research is now over 40 years old. Training technology continues to evolve rapidly, and it is appropriate to speculate that learners may be evolving as well. As the Army moves towards greater use of distributed learning technologies and gaming and simulations as specific training techniques, it is important to understand exactly how these delivery methods affect learning in current trainee populations, as well as how specific design features (e.g., background fidelity) might facilitate or hinder learning.

References

Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations: On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology, 89*, 92-103.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology, 90*, 414-434.

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93,* 187-198.

Sweller, J. (1994). Cognitive load theory, learning difficulty and instructional design. *Learning and Instruction*, *4*, 295–312.

Training for Efficient, Durable, and Flexible Performance in the Military

Contract #: DASW01-03-K-0002
Institution: University of Colorado

Contract Dates: 10/01/2002 to 9/30/2007
PI: Alice F. Healy
Co-PI: Lyle E. Bourne, Jr.
Liaisons: Stephen Goldberg, SSRU
Robert Pleban, IFRU

Problem(s)/Research Question(s) – This project aims to test and develop principles of training that promote efficient learning, durable memory, and flexible transfer performance.

Technical Barrier(s) – Learning, memory, and transfer performance are not always highly positively correlated; for example, rapid learning often leads to weak long-term retention. There is, thus, a technical barrier for finding conditions that simultaneously optimize all three aspects of training.

Significance/Impact for Basic Research – The experiments conducted as a part of this project have produced data that will support the development of a general theory of training encompassing learning efficiency, memory durability, and transfer flexibility.

Potential Transitions – Knowledge gained in this basic research may be usefully employed in the following ARI applied programs (among others):

- Virtual Intelligent Training for Objective Force Warriors (WP 294)
- FUTURE-TRAIN: Techniques and Tools for C4ISR Training of Future Brigade (WP 211)
- Future Force Warrior Training (WP 215)
- VICTOR: Virtual Individual and Collective Training for Objective Force Warrior (WP 233)
- Training for Interactive Distributed Environments (WP 214)
- Simulation-focused Collective Aircrew Training (WP 231)

Overview

We have developed a set of training principles that optimize the efficiency and durability of trained performance. In addition, we have discovered that conditions that lead to durability often, and perhaps always, lead to limited flexibility or adaptivity. In fact, we have found that training might have little or no benefit if there are discernable differences between the training and testing situations in the background or context, even if there are no changes made in the primary task requirements. The focus of the current project is, thus, to develop training procedures for knowledge and skill that will survive primary task or background changes and, thereby, produce flexible, as well as efficient and durable, performance in military tasks.

Research Approach

The experiments in this project are divided into three major groups. Experiments in the first group are designed to understand how individuals can be trained to contend with an unpredictable flow of information often large in quantity, rapidly presented, and ambiguous. Experiments in the second group are aimed to identify training factors that promote adaptive and flexible performance in the field. The final group of experiments examines performance in dynamic and changing task environments. We conclude with an effort to create, still in the laboratory, a complex set of tasks similar to those encountered by a digitally proficient pilot operating a fully computerized cockpit, and not unlike those of the digitally proficient "land warrior" (i.e., the Soldier of tomorrow). The major aim of this set of experiments is to determine the extent to which training principles, first established in simpler laboratory tasks, generalize to performance under these more complex conditions.

Accomplishments

In this summary, we focus on accomplishments in two experiments. Both of these experiments relate to understanding how individuals can be trained to contend with an unpredictable flow of information often large in quantity, rapidly presented, and ambiguous. The first experiment also is relevant to the identification of training factors that promote adaptive and flexible performance in the field. These two experiments have helped us to test and to increase the scope of previously proposed training principles.

The first experiment used a message comprehension paradigm in which we examined subjects' ability to remember and follow navigation instructions. We used this task to investigate variations in training to determine ways to optimize the understanding of, memory for, and execution of navigation instructions. In particular, this effort tested three different training principles that had been proposed in earlier research using different tasks: (a) *difficulty of training* (any condition that causes difficulty during learning may facilitate later retention and transfer), (b) *specificity of training* (retention and transfer are depressed when conditions of learning differ from those during subsequent testing), and (c) *variability of practice* (variable practice conditions typically yield larger transfer effects compared with constant practice conditions).

To test the relative strength of these principles, subjects were trained in one of three conditions. During training trials in the *easy* condition subjects were given only short messages including from one to three commands, in the *difficult* condition they were given only long messages including from four to six commands, and in the *mixed* condition they were given messages of all six lengths. Testing occurred following a short delay filled with an irrelevant experiment. At testing, all subjects received messages of all six lengths. On the basis of the difficulty of training principle, subjects should do better at test overall with hard training than with easy training. On the basis of the specificity of training principle, subjects should do better at test on short message lengths with easy training than with hard training but should do better at test on long message lengths with hard training than with easy training. That is, there should be an interaction between training condition and message length. On the basis of the variability of

21

practice principle, subjects should do better at test with mixed training than with either easy or hard training.

We found a significant interaction of condition and message length, as depicted in Figure 1. The mixed group was best on all lengths but length 2, and the easy group was better than the hard group on the easy lengths (1-3, which they trained on), whereas the hard group was better than the easy group on the hard lengths (4-6, which they trained on). The results support the advantages of both specificity and variability of training but do not support the hypothesis that difficult training would lead to overall best performance at test. Thus, in terms of our principles, we found support for the principles of specificity of training and variability of practice but no support for the principle of difficulty of training in the message comprehension task.
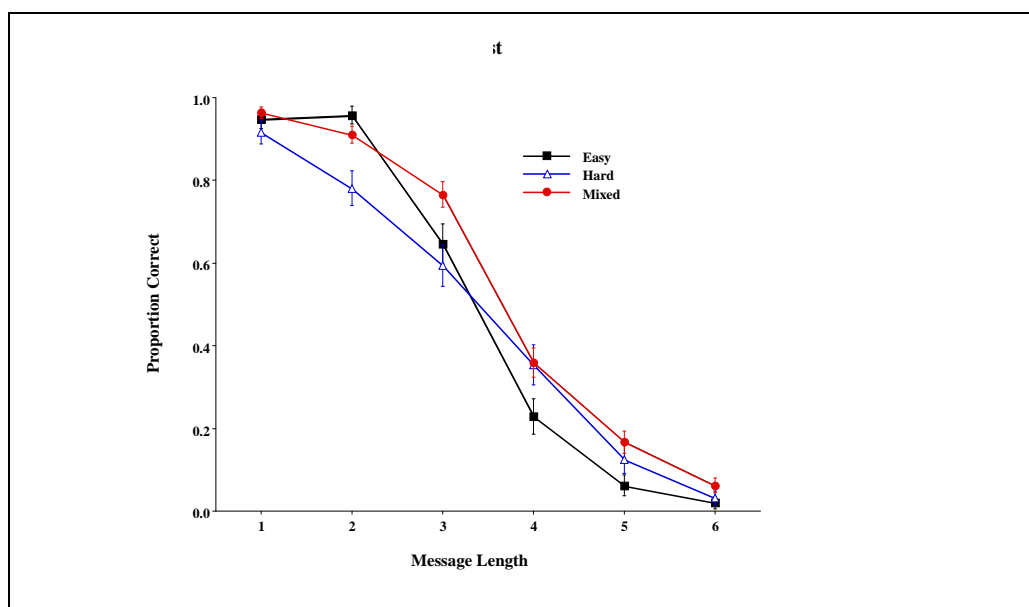
Figure 1. *Interaction of condition and message length.*

The second experiment continues a line of investigation in which we are attempting to understand the processes underlying the bow-shaped serial position function evident in studies of memory for order information. Earlier experiments in this series examined memory for information either learned outside the laboratory or learned in a single list presentation in the laboratory. In this experiment, we used a novel training procedure in which list items were presented multiple times. This procedure enabled us to compare the relative contributions of two variables postulated to affect the serial position functions: familiarity of the items and distinctiveness of the positions. Another novel aspect of the present effort was that it tested memory for spatial order, rather than memory for temporal sequence. Thus, unlike earlier studies, position distinctiveness was not confounded with temporal distinctiveness. Many military tasks require the participants to learn a spatial ordering that changes dynamically in time. For example, command personnel must learn where their own troops, enemy troops, and various combat resources are located on the battlefield. Furthermore, our manipulation of item

familiarity was much stronger than the manipulations used in previous studies of the serial position function.

We employed the technique we had developed earlier to investigate the strategic use of knowledge principle (learning and memory are facilitated whenever pre-existing knowledge can be employed as a mediator in the process of acquisition), in which subjects listed 10 of their friends and relatives, and these names of familiar individuals were included as stimuli in the experiment along with 10 names of unfamiliar individuals. There were two familiarity order groups that differed in terms of the assignment of names to list positions. In the familiar first group, familiar names were assigned to the first two list positions, unfamiliar names were assigned to the next two list positions, and so on throughout the list of 20 names. The opposite assignment of names to list positions was used for the unfamiliar first group. Following other procedures we developed earlier, subjects reconstructed the order of only a 12-name subset of the full list of 20 names. There were three serial position conditions that differed in terms of the 12-name subset of positions used for order reconstruction: 1-12, 5-16, and 9-20. As shown in Figure 2, there were striking effects of both familiarity order group and serial position condition on the shape of the serial position function, indicating that familiarity of the items and distinctiveness of the absolute list positions both contribute to the serial position function in reconstructing spatial order information. The effects of item familiarity were particularly striking, thereby also extending the strategic use of knowledge principle to tasks involving reconstruction of order. This research demonstrates clearly that prior knowledge can be used effectively to learn new position information, which should be relevant to training battlefield commanders who are required to learn location information, as described above.
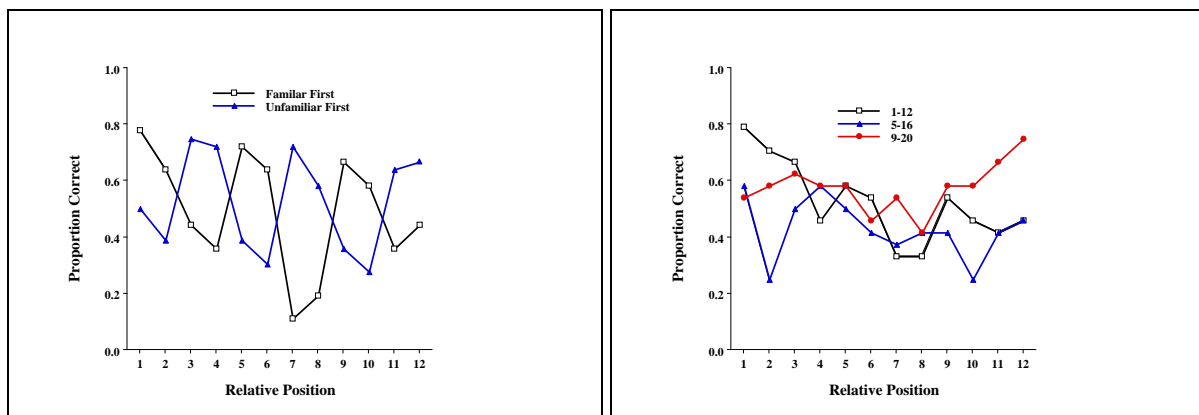
Figure 2 . *Effects of both familiarity order group and serial position.*

Contributions to Basic Science

We have proposed theoretical principles for optimizing training. The primary purpose of our experiments was to collect data that allow us to accept, reject, or refine each of these training principles that we formulated on the basis of our basic laboratory research and to develop and test new principles that are supported by these data. For example, our recent research of comprehension of navigation instructions allowed us to confirm the specificity of training and

variability of practice principles, and our recent research of serial order reconstruction allowed us to extend the strategic use of knowledge principle to a new domain.

Potential Army/Military Applications

Our experiments should be used to drive applied research. To illustrate this potential symbiosis between basic and applied research, we give two brief examples. First, our research has demonstrated a high degree of specificity from training to subsequent application. In fact, we have shown that training is specific even to the length of messages that need to be understood and executed. Test performance was best following training with all possible message lengths. These findings have crucial implications for military training because instructors may assume that teaching a particular task through a limited number of examples will generalize fully to an entire domain even when the examples differ in a fundamental respect (e.g., length) from the test situations. However, our findings imply that to be effective, training should incorporate a full range of examples on critical task dimensions. In future research, we will test the boundary conditions of the specificity of training and variability of practice principles using a variety of laboratory tasks and will explore additional ways to overcome training specificity beyond increasing practice variability. Although the tasks used in our research are often components of military tasks and require digital proficiency, they are not the real military tasks currently being trained in the Army.

We hope that applied research units are interested in testing whether the training principles we have developed would apply to such real tasks and whether the methods we hope to develop for overcoming the problem of training specificity could be adapted to improve military training. To that end, we visited and have had discussions with Stephen Goldberg (Chief, Simulator Systems Research Unit) about future collaborative research on the applicability of our training principles to simulator training. Second, we have shown that prior knowledge can be used to enhance the learning of spatial position information. In the future, we will conduct research to explore the boundaries of effective prior knowledge utilization, but the use of prior knowledge needs to be examined in the context of real military tasks as well. We hope that applied research units will be willing and able to take on this examination.

Final Summary

In some of our experiments, as in the investigation of memory for order information summarized here, we have shown that prior learning can aid in new learning of declarative information (or facts). These dramatic benefits of prior learning for declarative information stand in contrast to the surprising lack of prior learning benefits for procedural information (or skill) that we found in other experiments, including the investigation of message comprehension summarized here. Together these research efforts support a working hypothesis that expands on the earlier principles and models of encoding specificity, transfer appropriate processing, procedural reinstatement, and identical elements. These principles and models can account for the effects of training specificity that were observed, but they cannot explain without expansion when generalizability will occur instead of specificity. According to an expanded working hypothesis that we have developed, there is specificity (limited transfer) for tasks based primarily on procedural information, or skill, whereas there is generality (robust transfer) for tasks based

primarily on declarative information, or facts. Thus, we propose that for skill learning, retention is high but transfer is low; in contrast, for fact learning, retention is low but transfer is high.

Based on our findings and this working hypothesis, we recommend to the Army that any procedural skills be trained in circumstances as close as possible to those encountered in the field. We also recommend that when teaching new facts, these facts should be related to prior knowledge whenever possible so that the prior knowledge can enhance learning through mediation.

With respect to future researchers in the field of experimental cognitive psychology, we recommend further tests of our working hypothesis. Researchers also should attempt to develop new ways, similar to variability of practice, to overcome the limited transfer of procedural information and new ways, similar to the use of mediated learning, to leverage the robust transfer for declarative information.

Likewise, with respect to future Army related research, we recommend tests of the applicability to real military tasks of the working hypothesis contrasting procedural and declarative information in terms of their training, retention, and transfer. In addition, future Army related research should test the applicability to real military tasks of the ways we have developed in the laboratory both to overcome the limited transfer of procedural information and to leverage the robust transfer of declarative information.

The Integration of Implicit and Explicit Knowledge in Skill Learning

Contract #: W74V8H-04-K-0002
Institution: Rensselaer Polytechnic Institute

Contract Dates: 09/01/2005 to 10/31/2008
PI: Ron Sun
Co-PI: Robert Mathews
Liaison: Dan Horn, BRU

Problem(s)/Research Question(s) – How is implicit knowledge acquired from experience integrated with explicit knowledge from analytical thinking, and what role does this integration play in skill learning and in leading up to skilled performance?

Technical Barrier(s) – The interaction of these two types of knowledge is still poorly understood and more importantly, the integration of these two types of knowledge is heretofore unexplored. The process of interaction and integration appears to be highly complex, varied, and multi-faceted.

Significance/Impact for Basic Research – The major contribution of this work lies in coming up with a significant new theory that explains a range of human data (skill learning phenomena) on the basis of the integration of implicit and explicit knowledge. The integration of these two types of knowledge is heretofore unexplored in the literature. Therefore, this project will lead to advancing the psychology of skill learning and better understanding of this essential aspect of human cognition.

Potential transitions – Knowledge gained in this basic research may be usefully employed in:

- The Future Force Warrior Training program (WP 215)—Training programs that facilitate rapid accurate decision-making; and
- Various other training programs—VICTOR: Virtual Individual and Collective Training for Future Warriors (WP 233); SIMAIR: Simulation-Focused Collective Aircrew Training (WP 231); and Training for Interactive Distributed Environments (WP 214).

How implicit and explicit training interact and affect skill learning should be relevant to all of these programs.

Overview

It has been shown that high levels of skills depend on both conceptual (explicit) and subconceptual (implicit) knowledge. In our prior work, we have demonstrated that there is often a lack of awareness of experiential knowledge and, in fact, implicit experiential knowledge may sometimes be incompatible with explicit conceptual knowledge. These problems create serious challenges for training and learning in the military as well as the civilian contexts. The purpose of this research program is to better understand the integration of these two types of knowledge in the course of skill learning and how their integration may lead to better skill learning and better skilled performance. Specifically, we will (1) investigate the integration of implicit knowledge (from experiential learning) and explicitly formulated conceptual knowledge in the

development of skills and expertise, (2) explore when and how conscious reflection and resulting conceptual knowledge may be integrated with implicit knowledge to enhance, rather than to interfere with, experiential learning, and (3) develop computational models and theories of human skill learning that capture the fine-grained details of the integration of these two types of learning, which may be of use in understanding and improving training, on the basis of our experimental results.

Research Approach

This project investigates the integration of implicit and explicit knowledge in skill learning through experimental, theoretical, and computational means. In part, this work develops further a computational cognitive architecture CLARION, especially in terms of capturing of process details of bottom-up learning and top-down learning. These two processes represent two hypothesized ways of integrating implicit and explicit knowledge. Equally important in this project are a set of human experiments to explore different methods of facilitating the integration in order to enhance skill acquisition.

The human experiments will use a more complex process control task than those typically used in the implicit learning literature that we refer to as a diagnostic task. The task involves several interrelated input and output variables. The task captures real-world situations in which an operator (e.g., a manager) is attempting to control a main target variable (e.g., field performance) but must also monitor several other variables to be sure they remain in acceptable levels (e.g., desire to re-enlist, or morale). It also corresponds to situations where sequences of decisions must be made (selecting one or more input variables to use in each case) before feedback begins arriving about earlier decisions. This is analogous to a manager trying different treatments to improve workers' performance and waiting to see results, then making the decision to continue current treatment or to try something new. It also permits study of the effects on learning of individual variability in treatment outcomes (e.g., not all workers respond alike to the same treatment). While the task is dynamic, it provides feedback by discrete time steps (i.e., after each round of decisions, feedback is provided on the effect of the treatment decisions in each individual case). This division of performance into discrete time steps provides the opportunity to reveal learning strategies and to view changes in the learning processes over time.

In addition, the task contains a special model-building screen to develop and/or to reveal participants' explicit hypotheses about the relationships among the variables. Participants are told that their goal is to figure out the relationships among input and output variables and correctly map these relationships on the model-building screen. Responses on this screen reveal participants' conscious (explicit) knowledge about strength and direction of relations among input and output variables and indicate their confidence in their responses. They will also indicate what they plan to investigate in the next sequence of task performance.

In various experiments, opportunities to switch from the diagnostic task screen (assigning treatments to individuals and obtaining feedback from the past decisions on all output variables) to the model building screen will be controlled. In some experiments, participants will perform the task at a fast pace (they must respond every six seconds) to build implicit knowledge of the task while limiting time available to develop conscious hypotheses to test. Then, in a second

phase of the experiment, they will be allowed access to both screens to study how they translate their implicit knowledge of the task into an explicit model of the task. In other experiments, participants will be encouraged to engage explicit learning processes to learn how to perform the task.

Participants will be required to formulate and test their hypotheses by switching back and forth between the model screen and the task screen. These experiments will provide us with snapshots of how people attempt to implicitly and explicitly learn the task. Subsequent experiments will test various types of integrated training such as providing initial (partially reliable) explicit knowledge about the relations among input and output variables. Other variables to be studied will include the presence or absence of hints or cognitive aids (e.g., a history screen that allows participants access to previous inputs and their effects on a given individual), and type of task goal—whether the task goal is performance-based (e.g., getting all scores into the target zones) versus discovery-oriented (e.g., testing the effects of all treatment combinations).

Accomplishments

This research program began in September of 2005, starting with a major development of software for experiments. We are now completing the first series of experiments using the new software.

We have already made an interesting discovery. When a treatment has a strong positive effect on the main target output variable but also has a negative "side effect" on one of the other output variables, participants are generally unaware of the side effect but they discount (reduce) the estimated positive effect of the treatment on the main target output variable. This discounting effect is in proportion to the size of the negative side effect and does not go away even when participants are instructed to pay special attention to the output variable that has the side effect. This effect appears to be unconscious (implicit) but influences conscious estimates of effect size on the main target output variable. We are now conducting additional experiments to see if this effect increases or diminishes when training emphasizes explicit (responding to the model screen) or implicit (just trying to control the output variables) learning.

During this time, we also wrote and published two articles. One article involves two artificial grammar experiments that demonstrate the importance of including both bottom-up and top-down learning mechanisms to adequately account for artificial grammar learning. Across two experiments, we trained participants using a string edit task and highlighted relatively low-level (letters), medium-level (chunks), or high-level (structural, i.e., grammar diagram) information to increase the efficiency of grammar acquisition. In both experiments, participants who had structural information available during training generated more highly accurate strings during a cued generation test than other conditions, with equivalent speed. Experiment 2 revealed that structural information only enhanced acquisition when relevant features were highlighted during the task using animation. We suggest that two critical components for producing enhanced performance from provided model-based (explicit) knowledge involve: (1) using the model (explicit knowledge) to acquire experience-based (more implicit) knowledge rather than using a representation of the model to generate responses, and (2) receiving that explicit knowledge

precisely when it is needed during training. Our results were most supportive of the notion that a combination of top-down (model-based, explicitly driven) and bottom-up (experience-based, implicitly driven) processing might be a particularly powerful way to increase the acquisition of expertise (see, e.g., Sun, 2002; Sun et al., 2005). However, our work also suggests integrating these types of knowledge is not easy. It is clear that simply providing information about the underlying structure of the grammar and exposing participants to exemplars during training is not sufficient.

The second article involves a dynamic control task. Results demonstrate the utility of providing explicit knowledge before practice on the task, and more importantly, suggest how this information improves learning. Results also show that learning in this manner (with explicit knowledge provided) can lead to "costs" such as slowed retrieval, and that this knowledge may not always transfer to new task situations as well as experientially-acquired (mostly implicit) knowledge. Our findings also question the assumption that participants always acquire a highly specific "lookup" table representation while implicitly learning this task. We provide an alternate view and discuss its implications for theories of skill learning.

Simulations of various combinations of explicit and implicit learning also have been conducted (based on CLARION), and we are in the process of analyzing the results and developing new theoretical frameworks.

Contributions to Basic Research

The major contribution of this work lies in coming up with a significant new theory that explains a range of human data (skill learning phenomena) on the basis of the integration of implicit and explicit knowledge. One product of this effort is a conceptual framework, which addresses the ways in which these two types of knowledge may be integrated to produce expertise. This theory (the CLARION model) suggests that performance can be controlled by either a subconceptual knowledge base (implicit mode) or a symbolic conceptual mental model (explicit mode). Implicit control is fast but prone to error, particularly in early stages of skill acquisition. Explicit control is more accurate but slow to apply and prone to loss by forgetting over a retention interval. However, their integration may lead to better learning and better performance.

A computational cognitive architecture, CLARION, which is substantially different from other existing cognitive architectures, will be further developed in this work to simulate and capture a wide range of quantitative data that are related to the integration of the two types of knowledge, based on the above ideas. This will help us to explain (and eventually to predict) training and learning processes. We will carry out simulations to generate new insights and interpretations that can further explicate the integration of implicit and explicit processes. The outcomes (data, models, and theories) will likely provide a new perspective on skill learning. Our models and theories will be useful in better understanding human skill learning theoretically as well as in helping to improve training processes in practice.

Potential Army/Military Applications

This research links to applied programs on training and learning. Specifically, our basic research centers on understanding learning and cognition in terms of the integration of implicit and explicit processes, and thus it involves testing and developing learning and training principles that may be used to enhance human learning. The goals of the basic research align with those of the applied programs. Both focus on training effectiveness. Hence, work from our basic research could inform and be applicable to work in the applied programs. For example, our work may offer training principles that enhance training effectiveness, which can then be incorporated into an automated tutoring system. There are of course many other ways in which the results of our basic research program could be employed in the applied programs' systems, tools, and other products.

We expect that Army applications of this research will be in the area of training and human performance. A basic problem for the Army is how to ensure that novices in a Military Occupational Specialty (MOS) move quickly to more advanced performance (and perhaps to expertise) as a result of their training. In addition, most training focuses on teaching conceptual (explicit) knowledge rather than setting up the opportunity for gaining substantial experiential (mostly implicit) knowledge. While this may be appropriate for some specialties, some other specialties involve working with complex systems that are better learned initially through extensive experiences than with lectures or textbook lessons. As service personnel gain in expertise, it may often then become appropriate to focus more on conceptual (explicit) knowledge. In this process, the integration of the two kinds of knowledge is expected to be highly important. Our research will lead to some initial guidelines about how to best facilitate knowledge integration and thus learning and performance.

Future Plans

The new human experiments that will be conducted use a more complex process control task than those typically used in the implicit learning literature. This task (the diagnostic task) has a direct analog to real world situations such as making managerial decisions or clinical practice in medicine. In these situations, learners must contend with multiple sources of variability (i.e., unreliable explicitly provided information, individual variability in treatments, and client turnover). This task will allow us to study effects of these natural sources of variability on implicit and explicit learning outcomes, and their integration. Our previous efforts have found implicit learning to be more tolerant of learning under noisy conditions. We also plan to study the types of errors learners make under these conditions and how to prevent them.

In addition, the diagnostic task contains a special model-building screen to develop and reveal explicit hypotheses about the relationships among the variables over time. In some experiments, participants will perform the task at a fast pace to build implicit knowledge of the task. Then, in a second phase of the experiment, they will be allowed access to both screens to study how they translate their implicit knowledge of the task into explicit knowledge. In other experiments, participants will be encouraged to engage explicit learning processes to learn how to perform the task. They will be required to formulate and test their hypotheses by switching back and forth from the model screen to the task screen. These experiments will provide us with

snapshots of how people attempt to implicitly and explicitly learn the task. Other experiments will test various ways of integrating implicit and explicit training. Other variables to be studied will include various hints to help build accurate models and whether task feedback is performance (error detection) focused versus discovery (creativity) focused.

Simultaneously, we will develop the basic outlines of computational models for capturing various methods for integrating implicit and explicit knowledge in skill learning, namely various bottom-up and top-down learning algorithms. We will then test them against human data to be generated from human experiments. We also will work on further cognitive architecture developments.

We will continue to prepare manuscripts of theoretical and/or experimental papers, arguing for and developing our unique approach.

References

Sun, R. (2002). *Duality of the mind*. Mahwah, NJ: Lawrence Erlbaum Associates.

Sun, R., Slusarz, P., & Terry, C. (2005). The interaction of the explicit and the implicit in skill learning: A dual-process approach. *Psychological Review, 112*(1), 159-192.

BRU RESEARCH OBJECTIVE #2: PROVIDE FUNDAMENTAL KNOWLEDGE TO IMPROVE LEADER AND TEAM PERFORMANCE

Research under this objective provides concepts and methods for accelerating leader development, understanding and developing leader adaptability and flexibility, and discovering and testing the basic cognitive principles that underlie effective leader-team performance. Understanding the dynamics of small group leadership in face-to-face and distributed team environments is critical to this research objective as well.

Self-Initiated Development of Leadership Capabilities:
Toward Establishing the Validity of Key Motivational Constructs and Assessment Tools

Contract #: W74V8H-05-K-0001   **Contract Dates:** 11/01/2004 to 10/31/2008
Institution: Georgia State University   **PI:** Todd J. Maurer
                 Liaison: Larry Laffitte, LDRU

Problem(s)/Research Question(s) – The problem addressed in this research is to identify and validate new individual difference characteristics that are relevant to motivating leadership development, as well as ways of measuring those differences in people.

Technical Barrier(s) – Little is known about predicting leadership development behavior because most research has been directed at predicting the performance of leaders.

Significance/Impact for Basic Research – New behavioral constructs are needed that will be useful in understanding, predicting, and ultimately enhancing leadership development. It is expected that hypothesized relationships of these leadership development constructs/measures with each other and with additional predictors will help to establish their validity and potential usefulness in this domain.

Potential Transitions – Knowledge gained in this basic research may be usefully employed in:

- LEADERSHIP: Developing leaders in a changing army (WP 103),
- Accelerating Leader Development (WP 269),
- Accelerating Leader Development (WP 285), and
- SELECT21: Selection, Classification and Performance Metrics for the Objective Force (WP 257); ATO: IV.HS. 2002.01.

Overview

The Army (and other organizations) must have leaders or potential leaders who continuously pursue the development of leadership skills and who are flexible and adaptable through their involvement in self-initiated development and learning experiences. The Army cannot rely solely on mandated training or learning experiences—it needs leaders or potential leaders to be motivated to pursue self-initiated development of leadership skills and characteristics. Total development must be motivated and initiated from within the person. But little is known about predicting leadership development behavior because most research has been directed at predicting the performance of leaders. Leadership development and leader performance are not the same thing. Likewise, motivation to be a leader and motivation to develop leadership skills are not the same. The Army needs a better understanding of leadership development and those behavioral constructs and assessment methods that are relevant to leadership development, especially those that involve individual sustained effort to develop leadership attributes. Basic research that addresses these issues can ultimately lead to applied research that enables better selection and development of leaders.

A key first step in this process is to develop behavioral variables/constructs and assessment methods, including a network of related variables, that would help establish the validity and relevance of these new variables in this domain. The variables will be explored within a system of related constructs, including antecedents and consequences, to establish the relevance to other variables within the leadership development domain.

There will be two general sets of variables examined in this research. The first set of variables includes variables being introduced here as new to the leadership domain. This set includes: motivation for leadership development, self-efficacy (or self-confidence) for leadership development, and beliefs about the improvability of leadership skills/attributes (e.g., beliefs that leadership can be developed vs. leadership is born). In addition, participation in leadership development activities for the purpose of learning/developing leadership skills will be examined. The second set of variables to be examined are traditional personality variables, as well as selected variables from the employee development literature, that can be examined as predictors of these new variables. In general, it is expected that hypothesized relationships of these leadership development variables with each other and with additional predictors will help to establish their validity and relevance in the leadership development domain. Constructs relevant to leadership development (e.g., self-efficacy for leadership development, motivation for leadership development) will be distinguished theoretically and empirically from constructs relevant to leadership performance (e.g., self-efficacy for leadership, motivation to lead). Specific hypotheses and findings thus far are discussed below.

Research Approach

A two-wave survey method will be used to collect data from people. The source of subjects will be an on-line research pool through StudyResponse.com. As introduced at the 2003 annual meeting of the Society for Industrial and Organizational Psychology, StudyResponse is a service that matches researchers with participants willing to receive solicitations to complete surveys. Registered participants (currently N > 46,000) have varied demographic characteristics including racial makeup, gender, education, and a varied list of occupations. Working adults of a wide range of ages will be included, which will allow the investigation to focus on both younger and more mature adults. The relationships among the leadership development constructs will be tested when data for the total set of these measures is collected in two time periods. Collecting data for these measures twice allows for a retest interval to examine the stability of the constructs across time and also for prospective prediction of some of the constructs onto others at a later time.

Administrators of this service have found that results from the StudyResponse panelists' responses corresponded with results obtained in a national poll of opinions within just a few percentage points of error on just about every question asked. This suggests that the results provided by this sampling approach are reasonably representative of data provided by traditional and systematic sampling/polling procedures. This Internet panel approach provides a viable approach as a sound sampling strategy.

Accomplishments

In the first year (year one of a 3-year project), the literature was reviewed to determine the content of the newly prepared leadership development measures. Thirty-nine leadership attributes were identified. These attributes will be used to collect ratings for leadership improvability beliefs, self-efficacy for leadership development, and motivation for leadership development. Corresponding and more generally-worded measures of leadership development/improvability (as opposed to the development/improvability of specific leadership attributes) also were created. These measures and the measures for the traditional/existing personality and employee development variables were then reviewed for wording, scaling, and formatting. Several iterations and revisions were done. A computerized/Web version of the measures was developed and reviewed, and the system was tested to ensure the accurate capture of responses from the survey to the database. A pilot investigation was then conducted to test the surveys and the data collection process. Over 271 participants responded to all parts of this pilot investigation.

In year two, the data from the pilot were analyzed and the instruments examined to determine if any modifications needed to be made to the measures and Web survey prior to the full data collection. In this pilot test, in addition to testing the measures and data collection procedure, we obtained some preliminary findings that nicely set up the full research effort that was yet to come. In the pilot test, motivation to develop leadership was empirically distinguished from motivation to lead. Detailed statistical and psychometric analyses supported the validity of the new leadership development motivational measures so far. The leadership development constructs predicted leadership development activity reports better than the leadership measures. Analyses further suggested that being attracted to leadership for special rewards or extrinsic reasons is unrelated to the motivation to develop leadership capability, while being attracted to leadership for intrinsic or social responsibility reasons is significantly related to motivation to develop leadership. The results of the pilot test were presented at an international conference in April of 2007. Also in year two, after completing the pilot test analysis and determining that the measures and data collection procedure had integrity, we then initiated the full effort data collection. The full effort addressed predictors and outcomes of the motivational variables in a much fuller and detailed examination of the variables than in the pilot test, which was relatively limited and focused in nature. In the full research effort our major hypotheses were tested.

In year three, data collection was completed (909 respondents to T1 and T2 surveys) and the preliminary data analysis and interpretation revealed some very interesting and worthwhile findings. We have proceeded with proposing a new followup longitudinal data collection one year later; this will add new and valuable data, but also will require us to reanalyze our data to interpret the findings in light of the new sample(s). That task will be conducted in a fourth year, as soon as it is approved. Given the data we have so far, we can say preliminarily that it appears the results support the usefulness and validity of the measures. More specifically, the initial data analysis suggests that self-efficacy for leadership development (SELD) predicts motivation to develop leadership skills (MTDL), and that self-efficacy for leadership (SEL) predicts motivation to lead (MTL). The corresponding self-efficacy constructs have higher predictive validity for the corresponding motivational constructs (e.g. SELD→ MTDL; SEL → MTL). Also, leadership improvability beliefs (IBL) predict SELD, and IBL predicts SEL, although it

has a higher relationship with SELD due to their closer conceptual relationship. Further, MTL predicts MTDL, and this relationship appears to be surprisingly greater for those individuals who have lower (rather than higher) IBL, possibly due to a ceiling effect given that high IBL people have significantly higher MTDL on the scale. Also, as expected, MTDL predicts participation in leadership development activities moreso than MTL. While MTL does predict leadership development activities, it does not do so nearly as well as MTDL. It is likely that MTDL acts as a partial mediator between MTL and leadership development activities.

In terms of traditional personality variables, openness to experience has a relationship with both motivation to develop leadership skills and motivation to lead. Also, we found a relationship between extraversion and motivation to lead, and as expected it was higher than the relationship with MTDL. In addition, the traditional employee development constructs (e.g., self-perceptions that the person possesses the qualities needed for learning/developing, perceived need for development of one's own skills, social support for learning/developing, perceived benefits of developing skills) were related to both the development-oriented constructs (e.g., motivation to develop, self-efficacy for development), and the motivation to lead and self-efficacy for leadership constructs, but the overall pattern of relationships appears to be somewhat different. Overall, the results so far involving predictors of MTDL and MTL seem to confirm the different nature of the underlying MTL and MTDL variables, which was the purpose of these tests. We are exploring some new measures captured from supervisors of the particpants to get a sense for how the constructs predict their performance on leadership dimensions, potential as leaders and promotability, which is so far encouraging; the measures appear to have validity for predicting these variables. This will be explored more exensively when we collect followup longitudinal data.

Contributions to Basic Science

There has been a good bit of research on predicting leader performance. However, there has been relatively little research on predicting leadership development behavior. For example, while motivation to lead and self-efficacy for leadership have been explored in research, self-efficacy for leadership *development* and motivation to *develop* leadership skills has not. Also, especially given differing ideas about whether leaders are born or made, there may be differences in how aggressively and effectively potential leaders pursue the development of leadership skills. Successful development of leadership skills depends on strong motivation to develop these skills. The current research will test a model of leadership development motivation that specifically targets motivation to develop leadership skills.

A recent research article (Ruvolo, Petersen & LeBoeuf, 2004) authored by officers in the U.S. Army is titled "Leaders Are Made, Not Born," which strongly reflects the notion that leadership can be developed. Further, there are many examples of successful leadership training and development programs. However, it is well known that people differ markedly in their beliefs that leaders are *born* or *made*. Beliefs that leadership attributes are innate or fixed and are not acquirable or improvable might be a fundamental belief that can set the stage for someone's overall motivation for leadership development. To the extent that a person believes that leaders are born and not made, this can be a substantial impediment to optimal motivation for developing leadership skills. This research is exploring this belief and the role that the construct plays in

leadership development motivation. In addition, the similarities and differences between constructs associated with leadership performance and leadership development are being addressed conceptually and empirically. This includes the distinction between self-efficacy for leadership vs. self-efficacy for leadership development, and motivation to lead vs. motivation to develop leadership skills. Differences in the antecedents and outcomes of these constructs are being explored to establish their distinct construct validity. These issues have not been explored previously in relation to leadership development but do have great theoretical and practical promise.

Potential Army/Military Applications

Developing these new constructs in the leadership domain and establishing their construct validity in basic research could lead to applied research in key areas addressed by ARI. Generally, this applied research could fall into two categories: individual difference predictive studies and experimental research in which manipulations are used. First, these constructs and associated assessment methods, along with traditional predictors such as intelligence and personality, could be used as predictors of performance, motivation, development, and adaptability in Army training and field settings. To the extent that these measures predict effectiveness in adapting to leadership roles and learning and developing necessary skills, these tools might be used in selection and placement within the Army to identify those individuals who will likely most benefit from leadership training and who are most likely to take initiative to continually develop and improve leadership relevant characteristics in themselves. These individual-difference constructs and assessment tools might be explored in applied research for relationships with key behavioral and performance criteria relevant to leadership and leadership development and learning in practical Army settings.

There is a second way that developing these new constructs in the leadership domain in basic research could lead to applied research in key areas addressed by ARI. Research might be done on ways of influencing these constructs in Army personnel to improve motivation for leadership development. Variables such as self-efficacy are subject to enhancement through mastery experiences, persuasion, modeling, or vicarious learning. Self-efficacy for leadership development might also be influenced through these types of tactics. For example, some research has manipulated expectations for military personnel and found effects for increasing self-efficacy. Applied research might be conducted to determine whether self-efficacy for leadership development can be influenced, and whether this enhances motivation for development of leadership skills and subsequent development and performance. Prior research and theory suggests this should be possible.

It also should be possible to conduct applied research on ways in which beliefs about improvability of leadership skills can be influenced. Prior research in settings not involving leadership and leadership development suggests that implicit theories of ability (beliefs about improvability of attributes) are susceptible to influence. It should be possible to determine if influences on beliefs about improvability of leadership skills have a subsequent influence on motivation to develop leadership skills and attributes as people increasingly believe that it is possible to develop and improve specific characteristics. Thus, to the extent that these constructs are developed and validated in this basic research, additional applied research could follow.

Future Plans

In year 4, we plan to develop and collect longitudinal follow-up measures to include leadership development behavior/measures, leadership promotions, greater leadership responsibility received, and career advancement variables (such as salary, pay increases, bonuses, etc.) during the year. These data are important outcome variables that reflect on leadership and leadership development, and they would tell us the longitudinal predictive value of the measures we have developed for predicting these important behavioral and performance criteria. We could then answer questions such as: Do the measures actually predict leadership and leadership development outcomes at a later point in time? Can the measures predict who will be likely to get promoted into greater leadership roles in the future? Do the measures predict advancement in one's career as a leader, etc.? We will continue the full data analysis and interpretation, incorporating all new measures and sub-samples of respondents into the existing data, and examining interrelationships, as appropriate. We will complete the final interpretation of all data and write a technical report, integrating findings/data from all waves of data collection. Publication of this research in scientific outlet(s) will be pursued.

References

Ruvolo, C. M., Peterson, S. A., & LeBoeuf, J. N. G. (2004). Leaders are made, not born: The critical role of a developmental framework to facilitate an organizational culture of development. *Consulting Psychology Journal: Practice & Research, 56*(1), 10-19.

Language and Social Dynamics

Contract #: W91WAW-07-C-0029
Institution: University of Texas at Austin

Contract Dates: 05/21/2007 to 05/20/2009
PI: James W. Pennebaker

Problem(s)/Research Question(s) – What can we tell about the social dynamics of a working group based on the ways group members talk with one another?

Technical Barrier(s) – The measurement and interpretation of natural language in ongoing social groups is a new research focus. Linking group dynamics to the use of function words (e.g., pronouns, prepositions, and related words) has only recently become possible with the advancement of computerized text analytic tools.

Significance/Impact for Basic Research – How people use function words in everyday conversation provides information about how they are thinking and relating to other people. The analysis of language cues that people cannot readily manipulate (i.e., function words) provides a non-reactive and unobtrusive way to measure group dynamics.

Potential Transitions – Knowledge gained in this basic research may be usefully employed in:

- ARI and DoD management and groups training programs,
- The Intelligence Community's interest in remote sensing of groups and group dynamics, and
- Homeland Security's efforts at cultural interpretations.

Overview

Although communication through language is the primary way groups interact, virtually no one has studied it due to the limitations of language analysis technology. Only within the last decade have two overlapping technologies emerged that allow us to study language in groups efficiently: computerized text analysis programs and the Internet-based world of social communication. Historically, the analysis of written or spoken language has focused on the *content* of what is being conveyed. Now, in addition to the content of communication, it also is possible to explore the linguistic style that surrounds that content: that is, we are now focusing on people's use of function words, such as articles, auxiliary verbs, conjunctions, prepositions, and pronouns (Pennebaker, Mehl, & Niederhoffer, 2003).

Across multiple efforts, we are finding that the use of function words correlates with demographic factors (e.g., age, gender, social class), personality traits (e.g., self-esteem, neuroticism, conscientiousness), social factors (e.g., dominance, honesty, marital satisfaction, testosterone levels), and psychological state (e.g., depression, suicide-proneness). Function words are processed in the brain very differently than content words and are uniquely tied to basic social and psychological processes.

With recent advancements in computerized text analytic tools, we now are able to efficiently assess linguistic style within and across groups. By using these tools, we have started

to explore three new areas of research with groups that interact face-to-face (FTF) or by computer-mediated communication (CMC). The first broad research question asks how linguistic style can be used as a measure of social climate, including group task performance, liking, and leadership. The second research question has involved the examination of linguistic style and leadership in greater detail: that is, what is the linguistic style of a leader? What are the language cues that can predict who will emerge to be a leader in an initially leaderless group? What are the language styles of effective leaders? Our third area of research focuses on the continued development of text analytic tools and data mining techniques.

Research Approach

Our research uses a relatively non-intrusive methodology to assess individuals and groups in terms of cohesiveness, functioning, and individual-group relationships. We are continuing to develop an array of computerized text analytic programs to assess thousands of text samples along multiple dimensions. Our primary program, Linguistic Inquiry and Word Count (LIWC), (Pennebaker, Booth, & Francis, 2007), computes the percentage of words in a text that belong to a host of linguistic (e.g., articles, pronouns, conjunctions) and psychological categories (e.g., positive and negative emotions, cognitive mechanism words, social words). Using LIWC, we can compute how the words people use in natural language reflect personality and psychological states.

We are studying function word use in two ways. The first is to simply compare the actual rates of function word use by different people in a group or by different groups as a unit. For example, we can determine if successful groups use language differently from unsuccessful ones. The second is to examine how each person in a given group matches his or her function word rates with the other members of the group. This language synchrony measure is referred to as Language Style Matching, or LSM. Mathematically, LSM is based on how often a person uses each of nine function word categories relative to the average of their group. By isolating each person's ability to synchronize with others, LSM provides a unique look at social competence in a complex environment.

Beyond our lab experiments on group task performance, we are beginning to analyze group dynamics within written and transcribed oral transcripts that include emails, blogs, and overheard communications. Part of this research will focus on large natural language communities where we also have access to measures of social cohesiveness or conflict. Examples include working research teams, blog-writers linked to particular diseases, and even larger communities at the city level. By measuring word usage and LSM, we hope to evaluate the degree to which the groups work closely with one another and/or exhibit signs of tension.

Lastly, we are continuing to develop additional tools and algorithms to find meaningful patterns in natural language use, and for data mining natural language samples on the Internet. Part of this research will extend our text analysis tools to other languages such as Arabic.

Accomplishments

Language Style Matching in Groups. Although we have been funded only since the summer of 2007, we have run one experiment with 300 students who were assigned to same-gender groups of 4-6 people to work together on a problem solving task. Approximately half of the groups worked together in the same room (face-to-face, or FTF condition) and the remaining participants in the other groups worked in individual rooms using online chat technology (computer-mediated communication, or CMC group). The 25-minute task required participants to work together to come up with answers to each of 22 questions that required collaboration from all the members. In the FTF groups, all spoken language was recorded and transcribed; in the CMC groups, all online language was saved. The language from both groups was analyzed and compared with the outcomes of group cohesiveness and performance.

Groups whose members used a similar number of function words throughout the interaction (i.e., who showed high levels of LSM) were more likely to report enjoying the interaction and generally having been connected to the group. Although FTF groups evidenced somewhat higher degrees of matching and reported higher degrees of cohesiveness, both FTF and CMC groups employed verbal synchrony as an indicator of a cohesive group. Leaders had higher LSM scores than the group mean and than non-leaders. Despite this significant finding, leader LSM scores were not predictive of overall cohesiveness.
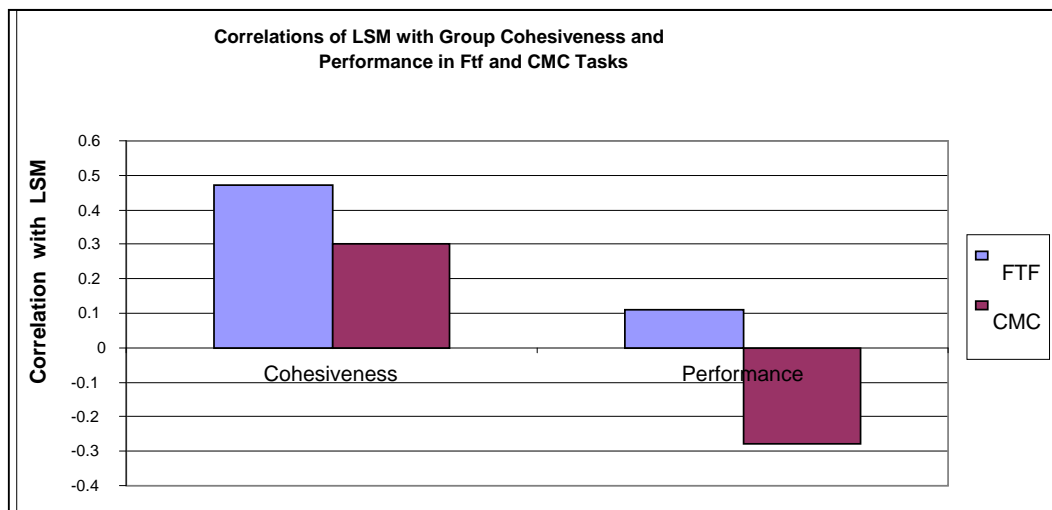
Figure 1. *Correlations of LSM with group Cohesiveness and Performance in FTF and CMC tasks.*

These findings suggest that while verbal synchrony is related to attraction in small work groups, it does not indicate that groups are working more effectively. There were no gender differences on LSM scores.

Language and Status. We also have been analyzing the results from a series of lab experiments that measure language and leadership. In the first experiment, same-gender pairs interacted on a FTF task, with one person designated as leader. The second experiment was similar but relied on a CMC form of interaction. A third experiment replicated the first two

experiments within the context of an informal "get to know you" setting where no formal task was involved. The final experiment analyzed hundreds of sent and received emails from 10 participants. The reliability of linguistic cues to status was assessed across all four experiments.
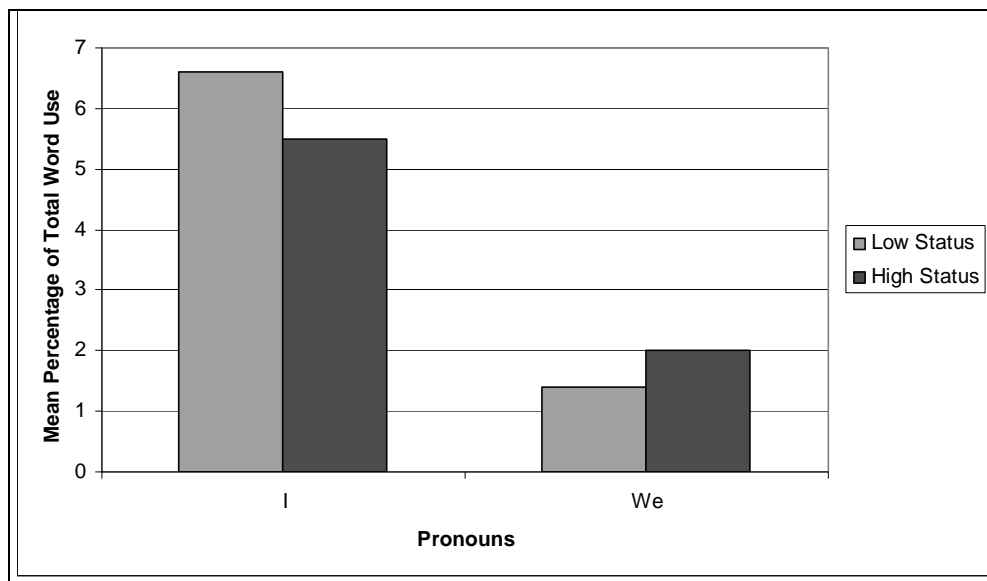
Figure 2. *Differential pronoun use by High Status and Low Status participants across FTF and CMC tasks.*

Word count, or how much each person spoke or wrote, was consistent across all four experiments. Higher status participants used more words than lower status participants. The most robust and impressive finding was the link between use of first-person singular pronoun and status. Lower status participants used more "*I*" across all four experiments. It is likely that lower status individuals are more self-focused and concerned with how a higher status person might perceive them. Higher status participants used first-person plural (e.g., *we, us, our*) at reliably higher rates relative to lower status participants, suggesting that higher status individuals were attempting to promote unity. Lower status individuals consistently used more negation words—words such as *no*, *not*, and *never*.

Similar effects have been found in the analyses of several of the Watergate tapes between Nixon and his aides (see Chung & Pennebaker, 2007). Those aides with the most egalitarian relationships with Nixon used I-words at comparable rates; those who were viewed as more subservient (Dean and Erlichman) used I-words at rates of over 50% above Nixon (see also Study 3 from Niederhoffer & Pennebaker, 2002). Overall, our results suggest there is a particular linguistic style associated with being a leader and with being a lower status individual across FTF and CMC interactions. These findings provide a baseline for identifying leaders and serve as a foundation for identifying the language of effective leaders.

Development of Text Analytic Tools. The analysis of function words provides important clues about the social relationship between a speaker and the audience, the psychological state of the speaker, and information about the speaker's social role and status in the community. Recently, we have begun to develop a computer strategy to measure and compare function word

use across several languages, including Arabic, Spanish, and German. We have started to develop an Arabic text analysis program based on our current computer program (named LIWC) that uses an English categorization scheme that would read Arabic text. We also have started to develop a LIWC program based on an Arabic categorization scheme that would process both Arabic and English texts. Although the development is in progress, we hope to give people who work with translators and translations an insight into dimensions of a culture that may be invisible to someone not familiar with the other language or culture.

Contributions to Basic Science

Language is generally thought of as a tool that we consciously manipulate. Changes in one's pitch or rate of speech happen outside of conscious awareness, but language is generally assumed to be a highly intentional communication channel. Our findings suggest that even the words we use are highly susceptible to influence by the affective nature of social dynamics; certain linguistic categories, such as function words, are as much a sign of how we feel as they are a tool for expressing it.

Through the development of our computerized text analytic tools, beginning with Arabic, we hope to uncover the function words that reflect social dynamics and psychological states across cultures in translated texts. Why should we care about the subtleties of translations? The translation of text from one language to another can convey the underlying pronouns and roles that it presupposes. The very act of translation reveals intended and implied meanings. It follows, then, that an examination of the function words that are dropped or added in a translation can provide explicit and measurable information to a culture's interpretation of a text. The development of computerized text analytic tools across languages can provide a reliable and novel measure of cultural differences in cognition and social dynamics.

Potential Army/Military Applications

The potential payoffs of this type of work are enormous for scientific advancement as well as for real-world problems faced by the Armed Forces and the population in general. Language analysis can be viewed as a form of remote sensing. Whereas we have always known that what people are saying is important, the primary contribution of this proposal is in noting that how they are conveying information also provides important information. With the advancement of computerized text analytic tools, we now have begun to systematically assess the language style of individuals within and across groups. We are beginning to see reliable language patterns across various experiments that we have conducted using small groups in tasks. Our approach may serve as a powerful way to assess leadership effectiveness. Indeed, we hope to evaluate if standardized natural language tests could predict and ultimately help to train leadership.

Based on previous efforts with our Spanish and German text analytic tools, we have begun to develop a text analytic tool for the analysis of language groups important to Homeland Security, such as Arabic. Counterintelligence Field Activity (CIFA) and other groups within and outside government constantly make significant decisions based on the translations of interviews, speeches, documents, and other verbal records. The net effect of such a computer strategy is that

we can provide translators, interpreters, or end users of translations with information about the social and psychological dimensions of any type of text sample. For example, the end user could receive the written text of a letter or speech, put it into our computer program in the original Arabic and receive output that could determine the degree to which it was conveying closeness, complexity, or whatever dimension that is afforded in Arabic. Similarly, one could analyze the translation of any text from English into Arabic to determine how Arabic speakers are likely to interpret the tone of the communication. This is a first stab at this way of thinking. If successful, the basic idea can easily be transported into virtually any language. The technology underlying this is transparent and the final computer programs are fast and easy to run. This project is both theoretically compelling and has an immediate and real world application.

Future Plans

Our plans for the next year of funding are to apply the findings from our lab to real world communities. We already have begun extending our collaboration with data mining experts to automate the downloading of online communities. For example, we are examining the language of entire blog communities to track responses, participation, matching, and the overall development and dissolution of social networks. In addition, we are aiming to begin transcribing archival data from real world groups working on complex tasks, such as flight simulation crews. By building a corpus of political speeches and interviews in both English and Arabic, we hope to be able to pilot our Arabic LIWC software. While our research aims to uncover group dynamics through language use, we are continually developing our data mining and text analysis toolkits that may be used for basic and applied research across the small groups domain.

<div align="center">

References

</div>

Chung, C. K., & Pennebaker, J. W. (2007). The psychological function of function words. In K. Fiedler (Ed.), *Frontiers of social psychology.* New York: Psychology Press.

Niederhoffer, K. G., & Pennebaker, J. W. (2002). Linguistic style matching in social interaction. *Journal of Language and Social Psychology, 21*, 337-360.

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). Linguistic Inquiry and Word Count (LIWC2007): A text analysis program. Austin, TX: LIWC.net.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology, 54,* 547-577.

Identifying Individual Attributes and Learning Principles That Foster Adaptive Performance and Promote Rapid Adaptability Skill Acquisition Through Multiple Modes of Development

Contract #: W74V8H-05-K-0004
Institution: George Mason University

Contract Dates: 01/15/2005 to 05/31/2008
PI: Stephen J. Zaccaro
Liaison: Robert Pleban, IFRU

Problem(s)/Research Questions(s) – What training and development principles produce the most effective acquisition and growth of leader adaptability skills? What individual attributes moderate the impact of training and development interventions on adaptability skills, and what combinations of leader attributes and team processes lead to adaptive performance in today's military operational environments?

Technical Barriers(s) – Past theories and models of leadership and team training are inadequate for understanding how to develop adaptive leaders and teams. Training and developing adaptability skills for both individual Soldiers and teams requires learning strategies that are different from those traditionally employed by the Army.

Significance/Impact for Basic Research – The successful completion of this research effort should contribute significantly to the literature on leader and team development, in particular on strategies for growing individual and team level adaptability skills through formal instruction, developmental work assignments, and self-development. We also should add new information to the small but growing literature on leader attributes patterns and their influence on leader adaptive performance.

Potential Transitions – The products of this research should provide the basis for applied research and application in various Army leadership training and development programs, such as those being conducted at Fort Benning and Fort Leavenworth. Other specific applications that can be explored through future potential DoD Small Business Innovation Research (SBIR) funding may include:

- The development of formal training programs that target leader adaptability skills, and
- The refinement and validation of assessment tools that can be used to measure developmental work experiences, learning attributes, and adaptive performance processes.

Overview

The overall purposes of this research effort are to expand current understanding about the nature of adaptive leadership and to explore the most effective ways of developing adaptability skills through three modes of training: formal instruction, operational assignments, and self development. We build on prior research to explore the leader attributes that promote effectiveness as environment conditions become more dynamic and adversarial. The specification of adaptive leader processes and attributes sets the stage for a consideration of best practices in leader training and development to foster growth in leader adaptability. Accordingly,

in this research effort we investigate training and development principles that are likely to produce the most effective growth of particular leader adaptability skills. We also examine how these principles vary across different modes of Army leader development.

We are extending prior research on adaptive performance to examine the influences of a training design that incorporates (a) a problem centered approach that utilizes multiple and qualitatively different performance practice scenarios, and (b) strategies for enhancing the practice of meta-cognitive thinking and performance regulation skills. The overall purpose of these training strategies is to enhance the degree to which trainees engage in greater frame-switching processes when operating environments become more dynamic. We also are examining and extending these training principles in the context of different training approaches, including work experiences and self-learning strategies.

This research effort also includes an examination of leader attributes that contribute to effective adaptive leadership. These personal attributes include such qualities as cognitive flexibility, self regulation and meta-cognitive skills, social and emotional intelligence, optimism, tolerance for ambiguity, and resilience. Relatively little research has examined how these attributes are integrated and how such integrations are linked to the processes of leader adaptation.

Research Approach

We are currently in the final year of this project. Earlier in this effort, we constructed an annotated bibliography on adaptability. We found 95 studies to date and coded them for the type of method used and the particular focus of analysis (adaptive performance, adaptability processes, adaptability attributes, training and development, and assessment tools). We also coded for level of analysis (individual, leader, and team). We found representative studies in all of our coding categories, with the fewest studies on the assessment of adaptability. We also found that additional research was most needed on:

- Predictive validation of linkages in models of adaptability attributes, processes, and performance;
- Validation of assessment tools; and
- Empirical examinations of training designs that focus specifically on the development of adaptability skills.

We continue to update this bibliography and will have a conceptual paper summarizing our full analysis of this literature at the end of the project.

In the first year of this effort, we completed an experiment that examined the effects of *experiential variety* in training practice scenarios on adaptive performance in teams. Adaptability typically requires a process of "frame-switching" that results in new and different ways of examining problems. Accordingly, adaptability training should provide opportunities to practice such frame switching. Experiential variety in training involves presenting trainees with qualitative shifts in performance scenarios, which in turn require fundamental changes in existing tactical and strategic plans. Such changes often follow a significant shift in the problem

solving frames leaders are using in particular problem domains. In the first experiment, teams completed a decision-making task (SimCity4), in which they practiced solving a series of three problems. In one condition, groups received a series in which all three problems required similar decision strategies. In another condition, groups received a series in which each problem required a qualitatively different decision strategy. Then all groups were given an adaptive performance trial that required decision strategies different from all previous trials. Groups were free to make as many decisions as necessary to solve the key issues in the adaptive performance trial. Possible decision solutions in this last trial ranged from low to high in their requirements for frame-switching. This research approach, then, tests the proposed association between one key training parameter—practice scenario variability—and team adaptability.

Accomplishments

The data from Investigation 1 indicated support for the proposed hypothesis. As Figure 1 shows, the variable scenario practice condition resulted in greater numbers of overall decisions than the routine practice decisions, but the largest differences were on decisions reflecting greater frame-switching. The frequency of high frame-switching decisions also was significantly correlated with a key parameter of overall adaptive performance.
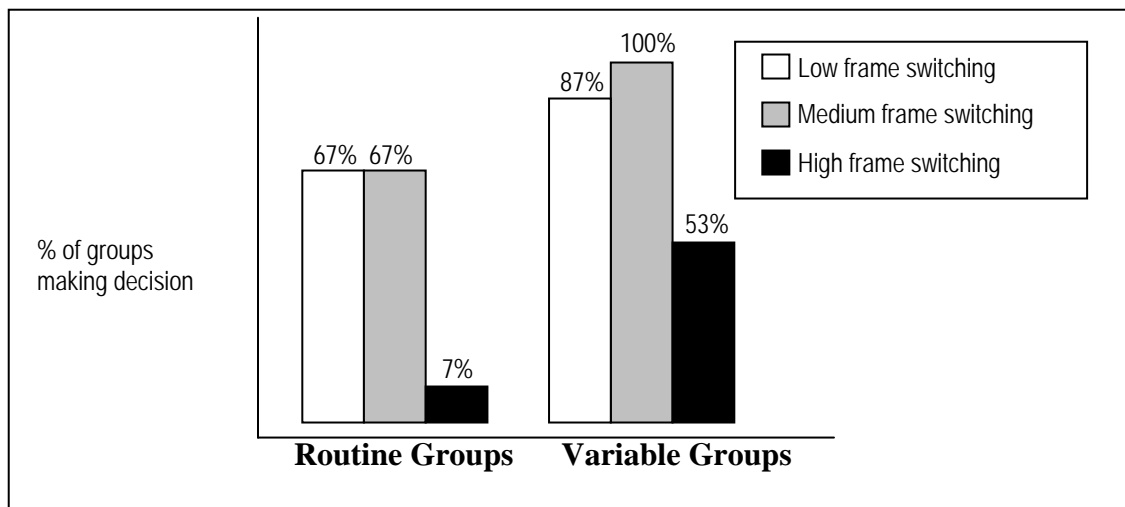
Figure 1. Percentage of groups making a decision across different group and frame switching conditions.

Note. N = 30. Groups make multiple decisions during the adaptive performance test trial. These decisions vary in terms of whether they reflect low, medium, or high frame-switching. This figure shows that the largest differences between the groups were in the number of high frame-switching decisions made by group members.

We assessed several potential mediators to explain the effects of training practice variability on adaptive performance, including collective efficacy for dealing with change, and the accuracy and sharedness of mental models. Groups did not differ significantly on these variables, suggesting that the primary consequence of practice variability is to prompt the occurrence of more frame-switching processes during adaptation.

In the second year of this effort, we completed a second experiment that examined the role of adaptive self-regulation training in fostering effective frame-switching. In Investigation 1, all groups had received such training before beginning the practice trials. In Investigation 2, all groups experienced variability in practice scenarios, but only half of them had received adaptive self-regulation training. In this experiment we found a pattern on adaptive decision making similar to Investigation 1: groups receiving qualitative variability in practice scenarios made greater numbers of decisions requiring higher frame-switching in the test trial than groups not receiving such variability. This effect was somewhat moderated by the receipt of self-regulation training. While this effect was a marginal one, it did indicate that groups receiving variable scenarios but no self-regulation training tended to make fewer decisions that reflected high levels of frame-switching. However, we found some interesting and unexpected interactions in analyses of overall performance in the adaptive performance trial. Such performance reflects all decisions made in the adaptive trial, not just those requiring high frame-switching. The effects of experiential variety on overall performance were moderated by self-regulation training and by the team's level of collective efficacy. An analysis of the interactions indicates that the effects of self-regulation training and collective efficacy on performance were stronger when experiential variety in training was low. Thus, each variable appeared to compensate for the lack of experiential variety in training, rather than enhance its effects. The interactions are shown in Figures 2 and 3.
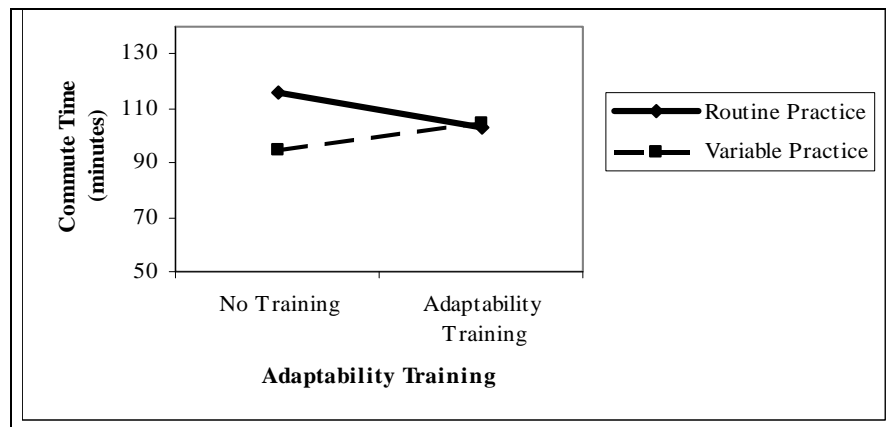
Figure 2. The interaction of training and practice on performance.

Note. Lower commute time equals better performance.

50
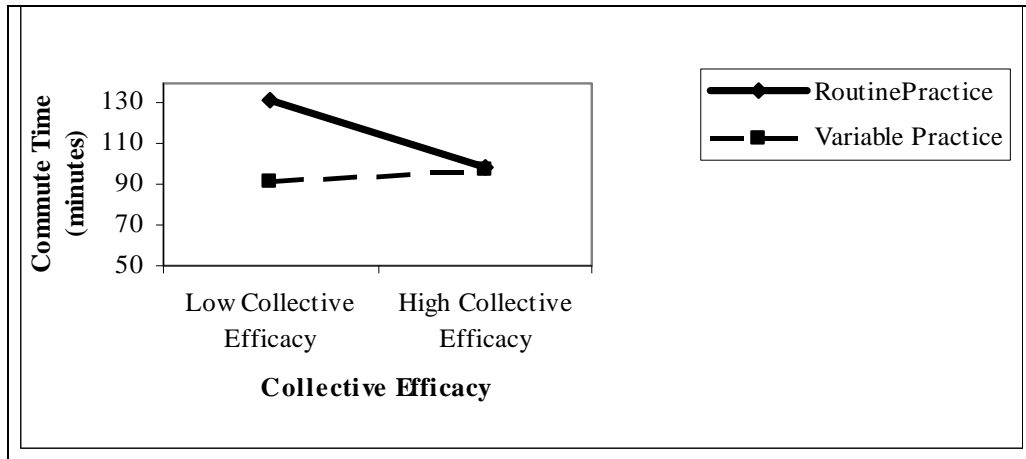
Figure 3. The interaction of collective efficacy and practice on performance.

Note. Lower commute time equals better performance.

We are completing two more experiments, one with college students and the other with Soldiers at Fort Benning and ROTC students to examine these effects more closely. Our focus is on exploring in more detail the changes occurring in individual cognitive processes as these individuals encounter the variable scenarios under high and low conditions of adaptive self-regulation training. The results of these efforts should shed some light on the role of both training parameters on cognitive processes linked to adaptive performance. We are also examining the role of individual differences in cognitive flexibility and meta-cognitive skills in moderating the proposed effects of training parameters.

In the second year of the project we also (1) completed construction of a survey that assesses qualities of work experiences proposed as contributing to growth in individual adaptability skills, and (2) completed the development of tools that can be used in self-learning programs focusing on leadership adaptability skills. We also developed a constructed-response assessment of adaptive decision making in complex and dynamic business scenarios. We are currently administering these instruments to managers in a variety of business settings to test the validity of these learning principles, as operationalized in developmental work assignments and self-development exercises. As part of these field research efforts, we are also administering measures of individual differences proposed as being associated with adaptability.

Contributions to Basic Science

This research effort should provide substantial benefits by enhancing our understanding of how to train and develop adaptability skills through formal course instruction, as well as through officer assignments and self-development efforts. This research effort should also contribute to the growing literature on how different characteristics of leaders are linked in patterns to influence leader adaptive performance, as well as the leader's responsiveness to different training interventions.

Potential Army/Military Applications

Some of the anticipated products of this research effort include training-based assessment tools and training protocols that can contribute to more effective development of Soldiers and officers. Several training units, including those at Ft. Benning and Ft. Leavenworth, are interested particularly in developing more adaptive leaders. We are using a computer-based training simulation at Fort Benning to test our hypotheses. We have completed a pilot round of data collection and are beginning to collect additional data.

Future Plans

We expect in the final year of this research effort to:

- Complete data collection and analysis of our third laboratory investigation with undergraduate students,
- Complete data collection and analysis for our investigation at Fort Benning and with ROTC students,
- Complete data collection and analysis of surveys to business managers that measure the quality of leader work experiences,
- Complete data collection and analysis on the self-development effort administered to direct marketing and sales managers, and
- Complete all data analyses and prepare research reports on each experiment, as well as a paper on the conceptual foundations of this research effort.

BRU RESEARCH OBJECTIVE #3: PROVIDE FUNDAMENTAL KNOWLEDGE FOR IDENTIFYING, ASSESSING, AND ASSIGNING QUALITY PERSONNEL FOR THE EVOLVING ARMY

Research in this section is directed toward identifying and measuring the aptitudes and skills that are unique to the human performance requirements of the Future Force. Exploring the sociological and psychological factors that could influence recruitment, retention, and Army performance are important aspects of this research objective.

Measuring Average Speed of Numerical Reasoning

Contract #: 1435-04-03-CT-74083
Institution: University of Minnesota,
              Twin Cities

Contract Dates: 10/01/2003 to 07/31/2006
PI: Mark L. Davison
Co-PI: Robert Semmes
Liaison: Michael Rumsey, SARU

Problem(s)/Research Question(s): Is time-constrained task performance within a given cognitive task domain determined by one ability, called *level* of performance, or by at least two abilities, level of performance and *speed* of performance?

Technical Barriers: (a) Undeveloped theory of the cognitive abilities responsible for time-constrained performance of tasks having non-trivial difficulty; and (b) the absence of valid models and procedures to measure speed of performance of cognitive tasks having non-trivial difficulty.

Significance/Impact for Basic Research: If our project is successful, we will have proposed the first psychometrically valid method for measuring average speed of performance of cognitive tasks having non-trivial difficulty. Moreover, the most widely accepted model of the structure of cognitive abilities (Carroll, 1993) will need to be revised.

Potential Transitions: This research could potentially lead to:

- Improved measures of key attributes that Soldiers need for effective performance in assignments for which mental quickness *and* accuracy are important,
- Enhanced predictive validity of the Armed Forces Qualification Test (AFQT) if that test were modified to measure performance speeds in each of the cognitive domains the test assesses,
- Performance scores that can be used in a model that matches persons to jobs (i.e., improved methods for classifying and assigning force members), and
- Improved training programs in areas where mental quickness and accuracy are critical.

Overview

Few would doubt that if each person in a group was assigned a set of non-trivial cognitive tasks of like kind, then people would vary in how successfully and rapidly they performed those tasks. Psychologists have known for several decades how to measure validly a person's ability to successfully perform a set of non-trivial tasks from a narrow cognitive domain (e.g., spatial visualization and inductive reasoning)—provided task performance is self-paced. Relative to a given cognitive task domain, that ability is called *level*.

On the other hand, there currently does not exist a valid way to measure how quickly a person successfully performs a set of non-trivial cognitive tasks when a strict time limit for completing each task is imposed. Few psychometricians appear to be aware of this methodological gap, or even aware that in a given cognitive domain a quickness or *speed* ability

might exist and be measurable. In part this may be because none of the several dozen (rarely cited) studies conducted in the 20th century to measure domain-specific speed abilities was successful.

Those previous attempts to measure speed abilities in specific cognitive domains failed primarily because investigators used faulty measurement models. For example, many investigators employed speed scores that were some function of a person's collective item response times (e.g., mean, median, or total). That approach, however, ignores the likely influences that varying item characteristics have on item response times, much as they do on binary item scores (right and wrong scores). Some researchers who used this approach also used both right and wrong answer times when computing speed scores, and thereby assigned speed scores confounded by extraneous variables such as times taken to guess answers and times taken to consider and then abandon items. Other researchers tried to adapt classical test theory's true-and-error score model to the measurement of speed. But that model—which was the dominant psychometric model for most of the 20th century—was, a priori, ill-suited for measuring a hypothetical speed ability. The true-and-error score model postulates that a person's success or failure on a cognitive test item, regardless of the conditions of performance, is a function of only one ability, the person's so-called *true score* in the given cognitive domain.

In this effort, we are using modern measurement models in an attempt to measure average speed ability in the domain of numerical reasoning. More formally, we are testing what we call the *speed-level hypothesis*: in narrow cognitive task domains, two abilities—speed and level—determine a person's success or failure in time-constrained task performance. Our null hypothesis is that level ability alone determines a person's success or failure in both time-constrained and self-paced task performance.

Research Approach

We have conducted a laboratory experiment wherein we have tested a large sample of University of Minnesota college students under two performance conditions: self-paced and time-constrained ("speeded").

Using numerical reasoning test items drawn from old, published editions of the SAT and GRE general tests, we constructed two equivalent test forms. Students were randomly assigned to take one or the other form in a self-paced test session, and then returned one or more days later and took the alternate test form in a speeded test session. This procedure in effect generated two, independent samples of students. In the speeded session, each test item bore an assigned time-limit. Each item's time-limit was set to the median correct response time computed for the given item using the self-paced test data of a large initial sample of students. In both the self-paced test session and the speeded test session, test items were administered one at a time using a personal computer. Each student's test sessions were conducted individually in the seclusion of a relatively sound-proof room. During each session computers recorded the student's answers and answer times.

To each of the two samples' self-paced test data we have fit a measurement model that postulates that a single ability, level, accounts for students' binary item scores (right or wrong

scores). The fitted models provide an estimate of each student's numerical reasoning level ability.

The centerpiece of our data analysis consists of our comparing the fit of two competing measurement models to our subjects' item scores earned during the speeded test sessions. One of the models, the level-only model, is the same as the one fitted to the self-paced data, and thus postulates that level ability alone accounts for the speeded item scores. The other measurement model postulates that two abilities—level and average speed—account for students' speeded item scores. Our speed-level hypothesis is confirmed provided the two-abilities model fits the speeded item scores better than does the level-only model.

 Accomplishments

- Based on our data analyses, the evidence for the speed-level hypothesis is mixed. In only one of our two subject samples have we found that item accuracy scores earned during time-limited task performance are unambiguously a function of two latent person attributes, one of which is level ability and the other of which almost certainly is an average speed ability. Our mixed results are likely attributable to two factors. First, our sample sizes compared to the number of test items we employed were too small for our analyses to detect an average speed ability. Second, our rule for setting item time-limits (a given item's time-limit was chosen as the median correct response time for the item obtained using the self-paced test data of an independent sample of students) resulted in time-limits for one of our samples that were too generous to elicit an average speed ability. To measure an average speed ability in numerical reasoning by modeling the binary item accuracy scores of persons who vary widely in level ability, test constructors evidently must set at least some item time limits *very* low—perhaps as low as the 10th percentile of the self-paced correct response times recorded for the item using an independent subject sample.

- To our surprise, in both of our subject samples we identified about a dozen students who correctly answered a higher percentage of items during speeded testing than during self-paced testing. Although there are other possible explanations for this phenomenon, perhaps for these students the challenge of solving numerical reasoning items bearing time-limits boosted their motivation to succeed over what they experienced during self-paced testing. The existence of such subgroups of people, although they are few in number, suggests that, at least in the task domain of numerical reasoning, relying solely on a person's level-ability score could well lead to erroneous conclusions about the person's actual capabilities.

- Modeling correct answer times, rather than binary item scores, may in the long-run prove a more fruitful approach to measuring speed ability in a given cognitive domain. Because answer times are measured on a continuous rather than a discrete scale, and because all who correctly answer a given item will inevitably vary in their answer times, correct answer times contain far more information about individual differences in speed of performance than do binary item scores. Using data from one of our samples, we successfully fit a speed-and-level model to 40 students' correct answer times to four

relatively easy items that all 40 students answered correctly. Using the fitted model, we assigned speed and level scores to each of the 40 subjects. Subsequently, we found that those students' speed scores, in linear combination with their level scores, performed better than level scores alone at predicting the percentage of 48 test items the subjects correctly answered during speeded testing. This finding is obviously consistent with our speed-level hypothesis. There is, however, a difficult technical barrier encountered when modeling correct answer times. If some item time-limits are short, every subject will miss at least some items during speeded testing. Speeded testing thus generates a varying number of missing correct answer times for each subject and each item. Because correct answer times are not missing at random, in order to reliably estimate each subject's speed ability, one must first statistically impute each subject's missing correct answer times. The methodology for such imputation is very new, and there is no assurance that a statistically defensible way to estimate the missing correct answer times can be found.

Contributions to Basic Research

In one of our two samples we have been able to confirm the speed-level hypothesis as applied to the numerical reasoning domain. We believe that our general approach to measuring average speed ability represents the first valid method of measuring (average) speed of performance in a cognitive task domain consisting of tasks of non-trivial difficulty. Our results suggest that, within a given task domain, the conditions of task performance alone can strongly influence what cognitive abilities account for performance. Our results further suggest that the structure of human cognitive abilities is likely much more complex than is portrayed by today's most widely accepted factor-analytic models (Carroll, 1993; Gustafsson, 1984). Provided our findings can be confirmed, and our measurement approach validated using much larger samples and a wider distribution of item time-limits, our method of measuring speed and level abilities will constitute a way to more completely and accurately summarize a given person's proficiency within the domain of numerical reasoning. Furthermore, should our findings be confirmed, we expect that our approach to measuring speed and level would likely transfer to many other cognitive task domains.

Potential Army/Military Applications

Quickness and accuracy in problem solving under emergency or combat conditions is a highly desirable attribute of any person serving in a command or key technical role in the Army. The desirability of this attribute increases in proportion to the pace at which technology increases the volume and speed of communications. However, none of the four subtests that comprise the AFQT provides a measure of performance speed in any of the domains the AFQT assesses. Rather, all four parts are intended to be tests of power (Russell, Peterson, Rosse, Hatten, McHenry, & Houston, 2001, p. 77). Having recruits' speed scores in addition to their level scores could lead to important improvements in matching recruits to job specializations, particularly in cases where those persons would need to be able to perform complex cognitive tasks not just successfully, but rapidly as well.

References

Carroll, J. B. (1993). *Human cognitive abilities*. New York: Cambridge University Press.

Gustafsson, J. E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence*, *8*, 179-203.

Russell, T. L., Peterson, N. G., Rosse, R. L., Hatten, J. T., McHenry, J. J., & Houston, J. S. (2001). The measurement of cognitive, perceptual, and psychomotor abilities. In J. P. Campbell & D. J. Knapp (Eds.) *Exploring the limits of personnel selection and classification.* Mahwah, NJ: Lawrence Erlbaum Associates.

Modeling the Strategic Role
of Basic Emotions with Cellular Automata

Contract #: W74V8H-05-P-0005   **Contract Dates:** 01/01/2005 to 04/31/2006
Institution: State University of New York,   **PI:** Craig DeLancey
           Oswego   **Liaison:** Bob Witmer, SSRU

Problem(s)/Research Question(s) – When are emotional behaviors (such as retribution in anger)—which in the short term may appear irrational—rational or beneficial for an individual or his/her group?

Technical Barrier(s) – We have no reliable way to induce, control, and measure anger, fear, and other emotions in controlled situations in human or animal subjects.

Significance/Impact for Basic Research – The research model, and a new explanatory framework, explain how some seemingly irrational emotional behaviors evolved and how they can be beneficial in some contexts. This provides predictions for human subject research.

Potential Transitions – Knowledge gained may be useful in studies including:

- ATO III.HS.2004.01 – Accelerating Leader Development (WP 269 & 285),
- LEADERSHIP: Developing Leaders in a Changing Army (WP 103), and
- Assessing and Developing Leaders for Special Operations Units (WP 218).

Overview

The goal of this research is to develop a model of the minimal conditions for certain kinds of emotional behaviors—particularly retribution in anger—to arise as a common strategy in a social group. The primary behavior studied was human performance in the Ultimatum Game, a widely researched kind of strategic interaction.

Emotions often motivate behaviors that are essential for our survival, but sometimes they can lead to behaviors that appear irrational. Some of these may in fact be irrational, but others may be behaviors that on balance benefit us if we understand them in the right context. We have little ability to distinguish these cases, or predict when they are likely to happen. Important tools for predicting and evaluating the decision making of individuals have been developed in economics and related fields of inquiry, but these tools are based on the assumption that individuals are rational agents who evaluate the expected outcome of their available choices, and then make choices that lead to their most immediate expected benefit. But actual empirical research into human decision making reveals that there are common kinds of behavior inconsistent with these assumptions.

A widely studied result of experimental economics is the seemingly irrational behavior of humans playing an Ultimatum Games (Roth, 1995). An ultimatum game is an anonymous one-shot two-player game in which one player (the "proposer") is given some utility (typically money) and must decide how to share this utility with another agent (the "receiver"). The

receiver can decide to "veto" the offer, resulting in neither individual getting anything. Traditional economic and game theoretic models predict that the receiver will never exercise veto, but will rather accept any offer greater than zero; and predict that the proposer, knowing this, will always distribute as little positive value as allowed in the game, and retain the rest. In actual experimental tests with human subjects, the veto is frequently used in response to low proposals, and proposers typically offer nearly even division of the utility. Thus, it appears that people often prefer less utility for themselves in order to punish others whom they perceive as making unfair distributions; and it appears that others know this and plan accordingly. One interpretation of this behavior is that individuals in the role of the receiver undertake retributive behaviors because they are angry.

The general notion that emotions have a special role to play in strategic action has received growing confirmation in a range of other sciences. However, progress in psychology and neural sciences fails to address two important features of the role of emotions in strategic action. First, existing models typically have no way to account for a benefit of emotional behaviors that stretches across games, or that influences the social context that in turn determines the structure of future games. For example, one utility of retributive behavior may be to enforce social norms (such as fairness in various distributions) that benefit on average the individual in question. Exploring such a possibility requires some modeling of social interactions over both a number of games, but also in a context where behavior shapes future interactions. Second, there is insufficient attention to the complex social interactions that such emotions may require. We may need to explain the role and utility of a basic emotion in an agent's action not only in terms of how it may have utility to the agent either immediately or by helping to maintain certain social norms from which the agent benefits, but also in terms of how the display of the emotion affects other agents.

Experimental conditions and existing models are not yet able to address such possibilities. Animal-subject experiments or human-subject experiments are inadequate because it is very difficult to induce fear, anger, and other emotions reliably and with controlled and significant intensities in subjects. It is even more difficult to induce such emotions in only selected subjects in groups. Study of the strategic role of basic emotions will remain intractable to human and animal experiments for some time to come.

Research Approach

The project used a technique called genetic algorithms to develop and test possible strategies that simple agents may use in a range of games or simple strategic situations. In this technique, strategies compete with each other, and how well a strategy does in comparison to other strategies determines how likely that strategy is to have "offspring" – that is, to be the model for future identical or slightly altered strategies. The goal was to determine whether it is sometimes beneficial to adopt retributive behaviors even when such behavior appears inconsistent with our standard notions of rationality. A series of hypotheses about the conditions for the evolution of such behaviors was tested against various games and various capabilities for the agents. Special attention was paid to the role of retribution in the Ultimatum Game, because the most extensive data on human subject performance is available for that game.

We also introduced a new form of explanation into evolutionary game theory. We considered the possibility that agents might confront different types of games that include similar kinds of interactions, and that an agent might develop a strategy for the most common game, and then apply this strategy to other similar games. This provides a compelling form of explanation for human performance in the Ultimatum Game.

Accomplishments

Our simulations confirmed our hypothesis that retributive behaviors that lack short term benefits can evolve (win out as strategies against other alternative strategies) in a population of interacting agents. Our primary model is a reputation-building model of altruistic retribution, which we call the "Reputation Game." In this model, simple agents interact with each other playing a game similar to the Ultimatum Game, but in which they can "know" the performance of other players in prior games.

Each agent has five parameters. Four parameters freely evolved (that is, the first generation had random values for each value for each individual, and then each generation is allowed to compete to have "offspring"). These parameters are: primary proposal, rejection-count threshold, secondary proposal (these three parameters apply when the agent is a proposer), and minimal acceptable proposal (the sole relevant evolving parameter when this agent is a receiver). The fifth parameter is the rejection count, which is a simple integer: this value begins as zero, and during a mock game play phase of thirty-two games as a receiver, each time an agent rejects an offer, this value is incremented. This can be thought of as the agent's reputation, announced at the beginning of a round, when the agent is playing the role of receiver. During the actual game play phase, games are played by each agent as a proposer, against some randomly selected other agents in the population. In each game, the proposer can see the rejection count of the receiver. If the receiver's rejection count is greater than the proposer's rejection-count threshold, the proposer will use its secondary proposal. Otherwise, it uses its primary proposal. This proposal is then communicated to the responding agent. When playing as a receiver, the agent will reject any offer not greater than or equal to that agent's minimal acceptable proposal value. After game play, the performance of an agent determines its fitness and thus the probability that that strategy, perhaps slight and randomly changed ("mutated") will be used in the next round of play by some agents.

This simulation evolves a stable strategy of both offering and expecting more equitable distributions in the Reputation Game. Figure 1 shows the mean population primary and secondary proposal values for 5,000 generations, normalized to range from 0 to 1. The typical population, after 300 generations, has both mean primary and secondary proposal values of 28% of utility. The population mean minimal acceptance value settles near 21% of utility.

Our hypothesis is that an Ultimatum Game could be seen as a kind of approximate instance of this Reputation Game. The Reputation Game is one shot but not anonymous. An anonymous interaction is an interaction where the proposer knows nothing of the receiver's rejection history. If we interpret this as the anonymous receiver having, from the perspective of the proposer, a rejection history of zero (or, alternatively, *any* rejection history), then the proposer will be offering a mean value around 9 out of 32. This is 28% of the utility. Camerer's

summary of Ultimatum Game findings shows a mean proposal typically ranging from 30% to 40% (Camerer, 2003). Furthermore, the population mean minimal acceptable proposal value tends towards 21% of the utility. Camerer's summary found that offers below 20% were rejected about half the time, whereas offers closer to 40% were rarely rejected. Thus, if we interpret Ultimatum Game performance as an application of the dominant Reputation Game strategy to the Ultimatum Game, these findings are close to human performance.
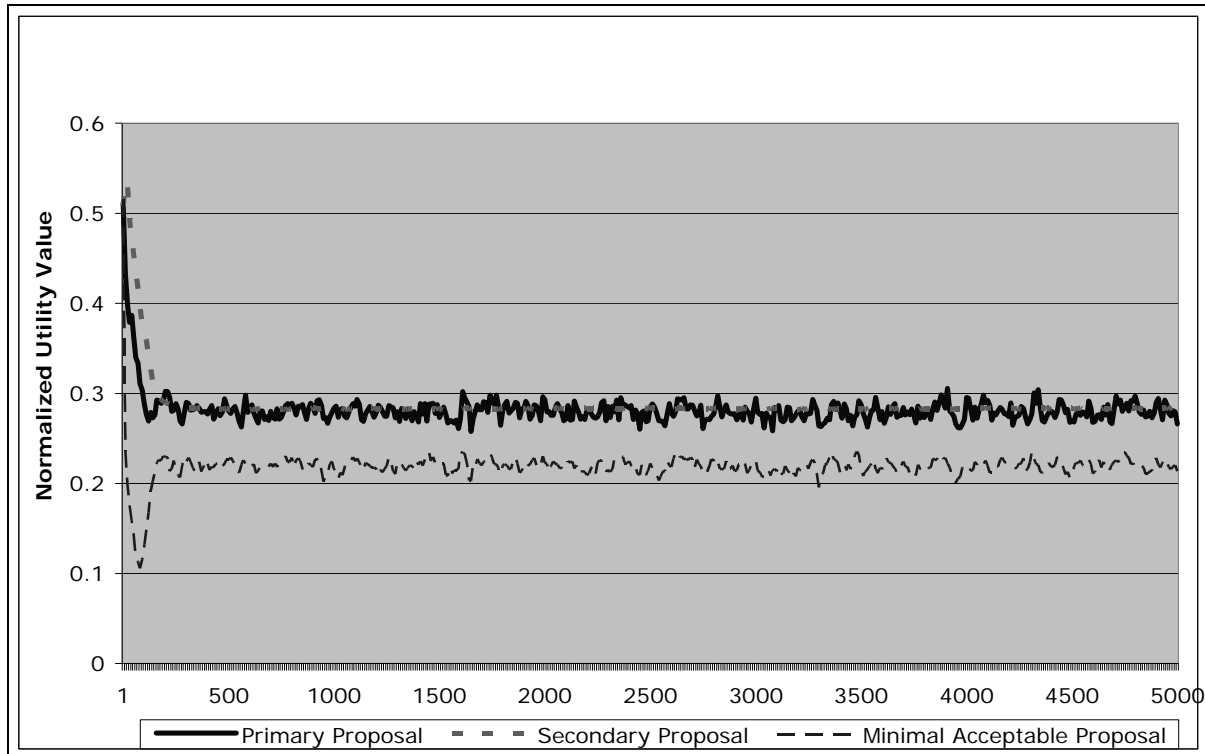
Figure 1. Population mean values for Reputation Game.

Contributions to Basic Science

The project provides a new model of altruistic retribution and extends standard game-theoretic or decision-theoretic approaches. The interpretation of the primary model suggests several new areas for future empirical human subject research. First, the model of retributive reputation building predicts that over very many anonymous interactions, less equitable distributions of goods would come to dominate. This could be tested empirically. Second, the model predicts that certain forms of rational control of emotion (those inconsistent with the primary inherited function of the emotions) will be difficult, and this is because of a biologically inherited resistance to reason that is beneficial to the individual. Finally, the model introduces a novel explanatory tool for evolutionary game theory: the interpretation of a game as an instance of a more common and general game.

Potential Army/Military Applications

These results suggest that anger and some other emotions must be recognized as inherited strategies in order to be effectively managed. These basic emotions resist rational control not because they are some kind of biological left-over which has not become amenable to our rational abilities—as if they were a kind of appendix of the mind. Rather, they resist rational control because it is part of their essential benefit to do so. This benefit is seen in at least two ways. First, short-term and seemingly irrational behaviors that signal and establish a reputation for retributive anger can be an important way to make credible threats. It is necessary, that is, for anger to resist reason in order for it to allow certain kinds of credible threats. The same is likely true of fear also. For example, fear could serve to allow an individual to make credible threats of defection even in the face of serious (deferred) penalties. Second, anger and fear may have a natural inherited tendency to resist rational control for the very reason that fitness and rationality can diverge. It might be irrational, given someone's beliefs, to avoid a dangerous situation (e.g., she may sincerely believe that some action is more important than the expected threat to her life the action will cause), but it may still be biologically fit for that person to avoid the action. Thus, for example, fear may in part resist rational control because this has been sufficiently fitness improving. Recognizing these facts can ensure that it is understood that these emotions are inevitable, and cannot be managed through direct rational argument. Rather, these are inherited strategic benefits that must be considered as universal elements of human social interaction between groups and between individuals within a group.

Final Summary

The results of our research extend strategic models of the kind found in game theory and classical economics to emotion-influenced behaviors that otherwise appear irrational. The model introduces a way to understand some of these behaviors, using the concept of strategies that are applied to similar but different situations. Thus, we can understand some perplexing and seemingly irrational instances of anger-motivated retribution—such as is often seen in one-time anonymous economic transactions, where retribution would seem to serve no purpose—as applications of tried and true strategies that the individual has found to be beneficial in the past and in more common kinds of situations that are neither one-time nor anonymous. For example, in the case of retributive anger, some retributive behaviors must be understood as a way to signal willingness to retribute within a context where reputation is essential to maintaining a social order. Furthermore, these retributive behaviors can be understood as behaviors that are naturally resistant to social control since they would be strategically ineffective if they were easy to defuse. For the Army, this may mean recognizing that emotional behaviors like these are inevitable, and cannot be managed through direct rational argument. Rather, these strategies arose because of inherited strategic benefits they yield in common forms of social interaction between groups and between individuals within a group.

Future research in this direction will include applications of the model as an explanation for norm-following (Bicchieri, 2005). One outstanding problem for social norms is how they arise in the first place. A plausible hypothesis is that some arise because of the over-application of a strategy. This research also may be of interest to those studying both the development of leadership and attempting to understand leadership. Leaders must understand the inevitable and

crucial role that emotions play in team dynamics, and also in reactions of Soldiers to stress and harm. Leadership training may benefit from an understanding and communication of the reasons emotions may be resistant to control.

References

Bicchieri, C. (2005). *The grammar of society: The nature and dynamics of social norms.* New York: Cambridge University Press.

Camerer, C. (2003). *Behavioral game theory.* Princeton, NJ: Princeton University Press.

Roth, A. E. (1995). Bargaining experiments. In J. Kagel and A. Roth (Eds.), *Handbook of experimental economics*. Princeton, NJ: Princeton University Press.

A Valid, Culture-Fair Test of Intelligence

Contract #: W74V8H05K006
Institution: Case Western Reserve University

Contract Dates: 11/15/2004 to 12/31/2006
PI: Joseph F. Fagan
Liaison: Peter Legree, SARU

Problem(s)/Research Question(s) – Can a racially unbiased test of the ability to process information predict how well young adults will succeed in college classes?

Technical Barrier(s) – Current theories of intelligence are based on an assumption that all those taking IQ tests have had equal opportunity for exposure to the information being tested. Thus, past efforts to develop an intelligence test that is culture-fair have not been successful.

Significance/Impact for Basic Research – The present effort provides further evidence to evaluate a theory that defines intelligence as information processing ability (Fagan, 1992, 2000). Specifically, how much you know (your IQ) depends not only on how well you process information (your intelligence) but also on the information given to you by your culture to process. Current research on a theory of intelligence as information processing finds racial differences in IQ to be due to cultural factors. A test of information processing promises to be the first valid, culture-fair test of intelligence.

Potential Transitions – The results of this research will aid in gaining the objectives of:

- The Army's Science and Technology Objectives (STO IV. SP. 2002.02 and DTO HS.52 and WP 111) as to Training of Future Combat Commanders and Staff,
- The Army's Science and Technology Objectives (STO IV. SP. 2003.06 and WP 215) for Future Force Warrior Training,
- The Army's Science and Technology Objectives (STO IV. SP. 2002.01 and DTO HS.54 and WP 104 and WP *257*) as to Selection, Classification and Performance Metrics for the Objective Force, and
- ARI's Work Programs on Leadership (WP 103, 211, 218, 269, and 285).

1. The general purpose of the present research is *to add fundamental knowledge to behavioral science… and to discover general principles …* [through an] *integrated programmatic effort to develop and to test theory* (as called for by RACO in BRO-BAA, page 1).

2. The specific goal is to discover the relationships among information processing ability and knowledge that lead to achievement and productivity in complex situations. The goal is in keeping with the Army's goal to improve its ability to *Select, classify, train, and/or develop Soldiers and leaders who …are adaptable…* [and who]…*can function effectively in …information rich…environments* (BRO-BAA, page 1, II, 1, a, b).

3. One of the BRO-BAA Basic Research Areas of Interest is Leadership skills. The proposed research will aid in the goal of identifying Leadership skills in Adaptability…when faced with novel situations (BRO-BAA, page 4, III, C, 1).

4. The research has resulted in a reliable, valid, and culture fair test of intelligence based on the ability to process new information. The research aids in further demonstrating that *intelligence is a function of an aptitude* [learning ability in novel situations] *that cuts across domains and how to develop a method for measuring this aptitude* [and to further] *validate* [this] *cognitive ability against relevant criteria*, i.e., academic achievement (BRO-BAA, page 5, III, D, 1).

In summary, the research is in keeping with the Army's goal to select Soldiers and leaders who can function effectively in information rich environments (BRO-BAA, page 1, II, 1, a, b) and to identify leadership skills as to adaptability when faced with novel situations (BRO-BAA, page 4, III, C, 1). This research will further aid in demonstrating that intelligence defined as learning ability in novel situations is an aptitude that crosses domains and in validating this cognitive ability against relevant criteria (BRO-BAA, page 5, III, D, 1).

Overview

The present effort is based on a theory (Fagan, 1992, 2000) that assumes that we act on the basis of what we know and that what we know is a result of our processing of the information provided to us by our culture. Based on these assumptions, the IQ score is seen a measure of knowledge. Specifically, how much you know (your IQ) depends on how well you process information (your intelligence) and also on the information given to you by your culture to process. The theory has been used to explain the source of IQ differences between Americans of different races. Specifically, we have found (Fagan & Holland, 2002, 2007) that American whites and African Americans who differ in IQ do not differ in their ability to process information when that information has been equally accessible to both groups.

Research Approach

College students are given multiple choice tests of their ability to acquire new information concerning the meanings of previously unknown words, sayings, similarities, and analogies. They also are tested for their knowledge of vocabulary in general, opposites, and analogies. An entire session (training and testing) lasts about 30 minutes. With the students' written consent, Scholastic Assessment Test scores, Grade Point Averages and numerical grades in a course also are obtained. Associations between performance on the culture-fair tests, academic aptitude (Scholastic Assessment Test scores), and specific achievement (cumulative GPAs and objective test scores in college courses) are analyzed.

Accomplishments

This research was undertaken to discover if a valid, culture-fair test of intelligence based on the ability to process new information could be developed. Our initial results, reported to ARI on 8/18/05, were based on a sample of 421 students, drawn from a private university, two

colleges and a two-year community college, who participated during the initial 10 month period of this research. We found that the tests of new learning, developed to that point, were culture-fair, reliable, and predictive of both academic achievement (numerical grades in class) and of a brief version of a standard test of scholastic aptitude (the Scholastic Assessment Test—Verbal).

Continued testing since 8/05, using a revised version of the test of new learning, has involved a higher percentage of community college students (relative to those in the private university and colleges) than was possible during the initial research period. A wider range of new learning, performance on the brief SAT-V, and academic achievement scores has allowed a more accurate determination of the role of information processing (new learning) and the role of the existing knowledge base (e.g., brief SAT-V scores) in the determination of academic achievement (grades). The results of the continued testing now are based on completed protocols from 596 students which, together with the sample of 421 students noted so far, yields a final total of 1,017 participants. (The planned sample size was 900 participants.) The latest results tell us, in accord with our previous efforts, that tests of new learning are culture-fair. African Americans perform as well as Whites. Further, the results demonstrate again that tests of new learning and tests of existing knowledge each contribute independent variance to the prediction of class grades, R (2/593) = .47, $P<.0001$, with Betas of .12 ($t = 2.4$, $P < .017$) and .39 ($t = 8.3$, $P < .0001$), for predictions from new learning and existing knowledge, respectively.

Two other findings are worthy of note. The first is that a subset of 18 of the 30 items tested on the tests of new learning proved to be particularly predictive of achievement, r (432) = .47, $P < .0001$, and were solved equally well by members of different racial groups. These 18 items may constitute a test of new learning to be employed in future studies where the inclusion of other measures in a protocol may be done while keeping session length within the temporal bounds comfortable for participants.

The second notable finding is that a brief, 24 item version of the SAT-V that we created turns out to be as predictive of college grades as the standard SAT-V. We constructed the brief SAT-V because our community college participants are not required to take either the SAT or the ACT for admission to college and we wanted a measure of knowledge that would be common to both community college students and university students. We also wanted a measure that would be comparable to the standard SAT-V as to validity. A sub-sample of 371 students attending the private universities provided the data for a comparison of our brief SAT-V and the standard SAT-V as to their comparative validity in predicting academic achievement. Measures of specific academic achievement (exam performance on objective tests) were obtained (with the students' permission) from the instructors of their courses. Cumulative Grade Point Averages representing at least two semesters of attendance also were obtained (with the students' permission) from the registrar at each institution.

Both the standard SAT-V and the brief SAT-V test were successful in predicting academic performance. Of most importance is the fact that the predictions from each test to academic performance were virtually identical. No significant difference was found between the SAT-V prediction of $r = .30$ to class grade or the brief SAT-V test's prediction of $r = .28$ to class grade, t (364) = 0.6. Similarly so for the predictions of each measure of aptitude to cumulative GPA ($r = .29$ and $r = .26$, respectively, with t (366) = 0.9. The results tell us that the scores

derived from our brief version of the type of questions typically asked on the SAT-V are as predictive of academic performance in college as are scores derived from the SAT-V itself.

Currently, questions have been raised in a series of articles in the *New York Times* (e.g., Lewin, 2005) about the amount of time necessary to complete the SAT, which is estimated at 3 hours and 45 minutes (and longer for students with disabilities). Earlier reports in the *New York Times* (e.g., Arenson, 2006) revealed a sizable error rate in the current scoring process for the SAT. In the present, ARI-sponsored effort, we found that a very brief, easily scored version of the SAT-V would be as predictive of college grades as the standard SAT-V.

In summary, we find there are at least two factors that determine success in a complex learning environment such as higher education: one is information processing (new learning) ability and the other is the extent of knowledge an individual has acquired to that point.

Contributions to Basic Science

Jensen (1998) believes that the IQ score reflects what he calls the general factor (or *g*) underlying intelligence: differences in IQ scores are due to differences in *g*. According to Jensen, genetic factors play a primary role in determining *g*. Thus, Jensen (2000) believes that it may not be possible to come up with tasks that show no differences in test performance between Blacks and Whites and yet still predict academic performance. Fagan (2000) does not agree with Jensen, pointing out that the failure to develop tests of intelligence that can be fairly applied across racial groups stems from a theoretical bias to equate the IQ score with intelligence rather than with knowledge. If intelligence is defined as information processing and the IQ score as knowledge, the possibility of culture-fair tests of intelligence based on estimates of information processing arises. Fagan is not alone in such a conjecture. Williams (2000, p.17) notes that " Fagan's ideas" (Fagan 2000) of measuring thinking or information processing rather than accumulated knowledge are relevant to the debate on intelligence testing and affirmative action because . . . a true measure of processing efficiency (if it could be devised) would be fair to members of all racial and ethnic groups." The chief theoretical implication of the results of the present research is that the assumption of equal opportunity for exposure to information made by theories of intelligence is false. When such an assumption is shown to be false, it becomes possible to develop culture-fair tests of intelligence that allow basic intellectual abilities to be recognized. Specifically, the inequality of educational achievement among races in our country has highlighted a need for culture-fair tests of intelligence. For example, Blacks do not do as well as Whites on IQ tests and other tests of knowledge such as the Scholastic Assessment Test. Basing admission to higher education on such test scores means that only a small percentage of Blacks are eligible for admission to colleges and universities. A culture-fair test of intelligence would allow basic abilities to be measured and would allow those with appropriate intellectual skills to pursue further schooling.

Potential Army/Military Applications

Two of the guiding missions of ARI, both historically and currently, are to provide new technology to meet the personnel challenges of the Army in recruiting, selecting and assigning quality people, and to provide scientific insight into leader development. The present research provides basic scientific data that may be applied by ARI to aid in *Soldier selection, assignment,*

and performance and in *leader development.* The results of the present research may lead to applied research that would be meaningful to the Army in the development of new, applied behavioral technologies to identify individual differences in cognitive ability that may be used in the selection of Soldiers and in the identification of mature, effective leaders. In addition, this research provides data to aid such selection and identification to be culture-fair with regard to gender and race. In general, tests of information processing ability that are valid predictors of achievement in complex situations will have important military application. Such tests can aid in providing a valid means to evaluate volunteers and select candidates for advanced education or training in complex situations of the sort performed by Army personnel. Such tests also can be employed in further selection and training stages. In the field, such skills are necessary to make quick and correct decisions based on rapidly incoming information. Skill in information processing is likely to be a key indicator of effective future Army leaders. Finally, using tests of information processing known to be culture-fair may provide an increase in the number of eligible recruits and an incentive to re-enlistment on the part of minorities.

Specific transitions to applied research include:

Selection, Classification, & Performance Metrics for the Future Force Soldier (STO IV. SP. 2002.1). Findings from our current research on the role of information processing ability can be applied to a personnel selection and classification system geared [toward identifying Soldiers that are] self-directed and effective information processors and problem solvers [a system that] might be incorporated into future enlisted selection and classification procedures [as well as being used as] outcome measures for future Soldier training and assessment.

Performance Measures for 21st Century Soldier Assessment. Our findings may be applied toward the development of a cost-effective…assessment system…for certifying a Soldier's qualification for job performance [and aid in the provision of] recommendations for an Army-wide system to certify the…qualifications of Soldiers for job performance.

Future-Train: Techniques and Tools for C4ISR Training of Future Brigade Combat Team Commanders and Staffs. The research focuses on how to measure cognitive skills that interact to influence performance in complex situations. Such aptitude may underlie the skills of future commanders and staff.

Leader Development, Accelerating Leader Development. The present research improves the assessment of cognitive skills, which will prove to be of importance to the selection and development of leaders.

Data on predictive validity relative to test length may be of benefit to the ARI applied program for Selection, Classification, & Performance Metrics for the Future Force Soldier (STO IV. SP. 2002.1).

Final Summary

Can a racially unbiased test of the ability to process information predict success in college classes? The present research is based on a theory (Fagan, 1992, 2000) that assumes that

how much you know (your IQ) depends on how well you process information (your intelligence) and also on the information given to you by your culture to process. Past research has found (Fagan & Holland 2002, 2007) found that American Whites and African Americans who differ in IQ do not differ in their ability to process information when that information is equally accessible to both groups. In the present experiments, college students were tested for their ability to acquire new information concerning the meanings of previously unknown words, sayings, similarities, and analogies. They also were tested for their knowledge of vocabulary, opposites, and analogies with a brief version of the SAT-V. Numerical grades in courses were obtained for these students. Tests of new learning proved to be culture-fair, reliable, and predictive of both grades in class and scores on the brief SAT. Tests of new learning and tests of existing knowledge (the brief SAT) each contributed independent variance to the prediction of class grades. A final notable finding is that the brief, 24 item version of the SAT-V created for this investigation is as predictive of college grades as the standard SAT-V.

References

Arenson, K. W. (2006). Testing errors prompt calls for oversight. *The New York Times*, 03/18/2006.

Fagan, J. F. (1992). Intelligence: A theoretical viewpoint. *Current Directions in Psychological Science*, *1*, 82-86.

Fagan, J. F. (2000). A theory of intelligence as processing: Implications for society. *Psychology, Public Policy, and Law*, *6*, 168-179.

Fagan, J. F., & Holland, C. R. (2002). Equal opportunity and racial differences in IQ. *Intelligence*, *30*, 361-387.

Fagan, J. F., & Holland, C. R. (2007). Racial equality in intelligence: Predictions from a theory of intelligence as processing. *Intelligence*, *35*, 319-334.

Jensen, A. R. (1998). *The g Factor*. Westport, CT: Praeger.

Jensen, A. R. (2000). Testing: The dilemma of group differences. *Psychology, Public Policy, and Law*, *6*, 121-127.

Lewin, T. (2005). SAT, at 3 hours and 45 minutes, draws criticism over its length. *New York Times,* 12/16/2005.

Williams, W. M. (2000). Perspectives on intelligence testing, affirmative action, and educational policy. *Psychology, Public Policy, and Law*, *6*, 5-19.

Achievement in Complex Learning Environments as a Function of Information Processing Ability, Knowledge, and Self-Control

Contract #: W91WAW-07-C-0026   **Contract Dates:** 05/07/2007 to 04/30/2009
Institution: Case Western Reserve University   **PI:** Joseph F. Fagan

Problem(s)/Research Question(s) – To what extent do information processing ability, knowledge, self-control, and race influence scholastic achievement and occupational success?

Technical Barrier(s) – Current measures of intelligence are estimates of knowledge, but do not control for opportunity for past exposure to the information upon which that knowledge is based. As a result, past efforts to develop culture-fair tests of intelligence that are predictive of scholastic or occupational achievement have not been successful.

Significance/Impact for Basic Research – The present research will provide additional evidence for a theory defining intelligence as information processing ability and for a method to assess intellectual ability in a culturally fair way. Exploration will be made of the manner and extent to which motivational and attitudinal factors interact with intelligence and knowledge to effect performance in complex situations.

Potential Transitions – The results of this research will aid in gaining the objectives of the Army's goal to select Soldiers and leaders who can function effectively in information rich environments (BRO-BAA, page 1, II, 1, a, b) and to identify leadership skills as to adaptability when faced with novel situations (BRO-BAA, page 4, III, C, 1). In addition, the proposed research will explore the applicability of principles of learning and self-development to mature adults in their 30s and 40s (BRO-BAA, page 5, III, C, 4). The proposed research will aid in demonstrating that intelligence defined as learning ability in novel situations is an aptitude that crosses domains, and in validating this cognitive ability against relevant criteria (BRO-BAA, page 5, III, D, 1). Finally, the proposed research will aid in understanding personal motivational factors, such as self-control, that are critical to success by operating together with cognition to produce achievement in difficult situations (BRO-BAA, page 5, III, D, 2 and page 6, III, F).

Overview

What determines achievement in a complex learning environment? Is it how well new information is processed? Is it how much relevant knowledge is brought to the situation? Is it a sense of self-control? Guided by theory, the purpose of the proposed effort is to discover how learning ability, knowledge, and self-control interact to determine achievement. Specifically, adults 18 to 24 years of age and adults 35 to 45 years of age, representative of the general population as to race and gender, will be tested for their ability to process new information, for the extent of their existing knowledge of word uses and meanings, and for their response to questions about their sense of self-control. These estimates of learning ability, knowledge, and self-control, respectively, will be employed in regression analyses along with the variables of gender and race to determine the extent to which each contributes independent variance to the prediction of scholastic achievement on the part of the younger sample, and to achieved occupational status on the part of the older sample.

Research Approach

Adults are given tests of knowledge of newly learned word meanings, sayings, similarities, and analogies to measure information processing ability. They also are given tests to measures their existing knowledge of word meanings, opposites, and analogies. All are asked to fill out a self-control survey and a questionnaire measuring their tendency to give socially desirable responses. Regression analyses will assess the variance contributed by information processing ability, knowledge base, self-control (minus social desirability), gender and race to grades for 18-24 year old college students and occupational success for 35-45 year olds.

Accomplishments

Since the inception of the effort (5/1/07), 347 young adults (18-24 years) have taken part in the research. Our findings to date are based on data from 122 of these young adults for whom all tests have been scored and entered and class grades (from Spring 07) are known. The results are straightforward. New learning ability, knowledge, and self-control are each predictive of class grades. An effort to examine the contributions of information processing ability, knowledge base and self-control to occupational success for 35-45 year olds also is well underway with 42 participants visited so far out of the 200 we plan to visit. Scoring of data from those tested to date has just begun. Based on our progress so far, it is anticipated that testing, scoring, and analysis of the data from the total planned sample (500 young adults and 200 adults in the age range of 35-45 years) will be accomplished within the time limits of this program.

Contributions to Basic Science

1. The general purpose of the effort is *to add fundamental knowledge to behavioral science…and to discover general principles* … [through an] *integrated programmatic effort to develop and to test theory* (as called for by RACO in BRO-BAA, page 1).
2. The specific goal is to discover the relationships among information processing ability, knowledge, and self control that lead to achievement and productivity in complex situations. The goal is in keeping with the Army's goal to improve its ability to *Select, classify, train, and/or develop Soldiers and leaders who…are adaptable…*[and who]…*can function effectively in …information rich…environments* (BRO-BAA, page 1, II, 1, a, b).
3. One of the BRO-BAA Basic Research Areas of Interest is *Leadership skills*. The research will aid in the goal of identifying *Leadership skills in Adaptability…when faced with novel situations* (BRO-BAA, page 4, III, C, 1).
4. A further area of interest that the research addresses is *Leader development.* In particular, *Maturity,* where the ARI BRO-BAA program calls for *developing a better understanding of the relationship between adult learning and growth in leadership ability.* The call is for *research to explore the applicability of …principles* [of learning and self-development] *to mature adults in their 30s and 40s* (BRO-BAA, page 5, III, C, 4).
5. The research is a continuation of our current ARI sponsored research in which reliable, valid, and culture fair tests of intelligence based on the ability to process new information were developed. The research will aid in further demonstrating that *intelligence is a function of an aptitude* [learning ability in novel situations] *that cuts across domains and*

how to develop a method for measuring this aptitude [and to further] *validate* [this] *cognitive ability against relevant criteria*, (i.e., academic achievement and occupational success) (BRO-BAA, page 5, III, D, 1).

6. The research will aid in understanding *personal motivational factors* [such as] *self-control...which are critical to success...*[by] *operating together* [with] *cognition... to achieve…in difficult situations* such as academic or occupational success, or, with regards to Soldiers, in *difficult military operations* (BRO-BAA, page 5, III, D, 2 and page 6, III, F).

Potential Army/Military Applications

Two of the guiding missions of ARI, both historically and currently, are to provide new technology to meet the personnel challenges of the Army in recruiting, selecting, and assigning quality people and to provide scientific insight into leader development. The goal of the present research is to provide basic scientific data that may be applied by ARI to aid in *Soldier selection, assignment, and performance* and in *leader development.* Specifically, the results of the research may lead to applied research that would be meaningful to the Army in the development of new, applied behavioral technologies to identify individual differences in cognitive ability, knowledge, and self-control, which may be used in the selection of Soldiers and in the identification of mature, effective leaders. In addition, the research will provide data to aid such selection and identification to be culture-fair with regard to gender and race. In general, tests of information processing ability, knowledge, and self-control that are valid predictors of achievement in complex situations will have important military applications. Such tests can aid in providing a valid means to evaluate volunteers and select candidates for advanced education or training in complex situations of the sort performed by Army personnel. Such tests also can be employed in further selection and training stages. In the field, such skills are necessary to make quick and correct decisions based on rapidly incoming information. Skill in information processing, knowledge, and self-control are likely to be key indicators of effective future Army leaders. Finally, using tests of information processing known to be culture-fair may provide an increase in the number of eligible recruits and an incentive to re-enlistment on the part of minorities.

Specific transitions to applied research include:

Selection, Classification, & Performance Metrics for the Future Force Soldier (STO IV. SP. 2002.1). Findings from the present research on the role of information processing ability, knowledge, and self-control in achievement can be applied to a personnel selection and classification system geared [toward identifying Soldiers that are] self-directed and effective information processors and problem solvers [a system that] might be incorporated into future enlisted selection and classification procedures [as well as being used as] outcome measures for future Soldier training and assessment.

Future Classification and Assignment. Findings from the research may aid in the development of a person-job match model [which incorporates] motivation as well as aptitudes.

Performance Measures for 21ˢᵗ Century Soldier Assessment. Findings from the research may be applied toward the development of a cost-effective…assessment system …for certifying a Soldier's qualification for job performance [and aid in the provision of] recommendations for an Army-wide system to certify the…qualifications of Soldiers for job performance.

Future-Train: Techniques and tools for C4ISR Training of Future Brigade Combat Team Commanders and Staffs. The research focuses on how to measure cognitive skills, knowledge and self-control which interact to influence performance in complex situations. Such aptitude and motivation may underlie the skills of future commanders and staffs.

Leader Development, Accelerating Leader Development. The proposed research may improve the assessment of cognitive skills, self-control, experience based learning, and the acquisition of tacit knowledge. Each of these factors may prove to be of importance to the selection and development of leaders.

Future Plans

As noted above, since the inception of this effort (5/1/07), 347 young adults (18-24 years) have taken part in the research and findings to date are based on data from 122 of these young adults for whom all tests have been scored and entered and class grades are known. Collection of data from 35-45 year olds also is well underway with 42 participants visited so far out of the 200 we plan to visit. Scoring of data from those tested so far has just begun. Based on our current progress the plan for the future is that testing, scoring, and analysis of the data from the current sample will be accomplished by the end of the first year of the research program (4/30/2008) and that testing, scoring and analysis of the data from the total planned sample (500 young adults and 200 adults in the age range of 35-45 years) will be accomplished along with preparation of associated reports and submission of manuscript(s) for publication within the remaining time limits of this research, that is from 11/17/07 until 4/30/2009.

Testing a Theory of the Determinants of Individual Job Performance for United States Army Junior Commissioned Officers

Problem(s)/Research Question(s) – How do individual difference variables work to influence performance on specific performance dimensions relevant to junior commissioned officers in the U.S. Army?

Technical Barrier(s) – There is no well-established theory of the process by which individual-difference variables predict citizenship performance or adaptive performance, in either a military or a civilian context.

Significance/Impact for Basic Research – Support for this theory of performance contributes to a better understanding of the relationships between predictor variables, mediating variables, and job performance criteria.

Potential Transitions – ARI has created various temperament scales thought to be related to adaptive performance and has empirically linked these measures to the mission performance of Special Forces (SF) Soldiers. ARI is now leveraging this measure to create a self-development feedback packet for SF Officers. The measures created in this basic research project are designed to predict performance at a more proximal level than the current temperament scales, so they could be used to revise or improve this self-development package or other Soldier self-development packages that are used outside of SF.

Overview

Organizational citizenship performance and adaptive performance are increasingly important components of the Army junior commissioned officer job performance construct domain. Despite the importance of these constructs, there has been very little research studying them in a military setting. There is also no well-established theory of the process by which individual-difference variables predict citizenship performance or adaptive performance, in either a military or a civilian context. In this research, we tested a theory of the process through which individual difference variables work to influence performance on specific citizenship and adaptive performance dimensions relevant to junior commissioned officers in the U. S. Army. Johnson (2003) proposed a theory of performance determinants that goes beyond earlier theories proposed by Campbell, McCloy, Oppler, and Sager (1993) and Motowidlo, Borman, and Schmit (1997), primarily by specifying an expanded conceptualization of motivation. Johnson broke the motivation component into specific motives (e.g., organizational commitment, job satisfaction), proactive cognitive processes (e.g., self-efficacy, goal commitment), and on-line cognitive processes (e.g., self-regulation). These variables, along with knowledge, skill, and habits, mediate the relationship between indirect performance determinants (e.g., cognitive ability, personality) and any kind of individual job performance construct.

Support for this theory would provide researchers and practitioners with a tool that can be used to identify the constructs through which individual difference variables work to influence performance on specific dimensions. This is especially important for identifying predictors of citizenship and adaptive performance. Such a tool would contribute to a better understanding of the relationships between predictors, mediators, and job performance criteria, and could be used to choose appropriate predictors for whatever criterion construct is of interest for a particular Army job. A rigorously developed and tested theory of the influence of individual differences on job performance would not only facilitate selection, classification, and training of junior commissioned Army officers, but would also add significantly to our evolving understanding of job performance.

Research Approach

In order to test this theory, we assembled and developed a battery of instruments that are construct-valid measures of each component of the theory. All measures created for this project were based on input from ROTC cadets. To measure task, citizenship, and adaptive performance, we created a multi-source performance rating instrument measuring performance on five different dimensions. We created a past behavior record to measure skill and a situational judgment test to measure knowledge, both of which are relevant to five performance dimensions. We also created (a) a unique self-report measure of work habits; (b) a motives scale assessing attitudes, values, and preferences that predict performance; and (c) a motivation scale measuring self-efficacy, expectancy, goal content, and goal commitment. We also created a self-regulation scale to measure the extent to which respondents were able to focus on performing the behaviors they indicated they intended to perform. We also assembled a battery of existing instruments to measure cognitive ability, personality, experience, and action control.

The predictor battery was administered to 155 ROTC cadets in the Fall of 2006. Criterion administration took place during March and April of 2007, allowing us to assess the extent to which cadets performed the behaviors they intended to perform. The self-regulation scale was administered four times during this period, allowing us to track how well each cadet was able to maintain focus on performing these behaviors. Each cadet was rated by one to five peers. These data were analyzed using structural equation modeling to test the Johnson (2003) theory, as well as several ancillary hypotheses.

Accomplishments

We found that modified versions of the Johnson (2003) model described the process by which predictor variables influence performance. Different models were necessary to describe the process for task, citizenship, and adaptive performance. The strongest support for the model came from the task performance dimension. For this dimension, knowledge, skill, cognitive choice aspects of motivation, and self-regulation were all direct determinants of performance, and each of these determinants was predicted by different combinations of personality traits, experience, and motives. Motives such as job satisfaction and affective commitment were related to performance only through the mediating influence of expectancies and goal commitment. A similar model was supported for the adaptive performance dimension. The differences were that

only some of the hypothesized direct determinants had significant paths to performance, and one personality trait had a direct influence on performance.

This model did not generalize to the citizenship performance dimensions, although a similar model was found for each different dimension of citizenship performance. In this model, motives were a direct determinant of performance and expectancies were not. Knowledge, skill, and self-regulation were direct determinants, but never all at once and the one that was a determinant depended on the citizenship dimension of interest. There were also two cases in which a hypothesized indirect performance determinant had a direct influence on performance.

There are several implications of these results. First, we have made great progress in identifying the mediators through which personality, ability, and experience operate to influence individual job performance. These individual differences lead to job performance because they influence the more proximal performance determinants of knowledge, skill, cognitive choice aspects of motivation, self-regulation, or motives. Second, there appear to be fundamental differences in how the model operates depending on the type of performance predicted. For example, the model that describes how individual differences influence task performance is different from the model that describes how individual differences influence citizenship performance in several ways. Motives such as job satisfaction directly influence citizenship performance but only indirectly influence task performance. There are also more direct determinants of task performance than of citizenship performance. Finally, the influence of motivation on performance is far more complex than suggested by previous models. There are several components of motivation and each component may be a direct determinant, an indirect determinant, a mediator, or a suppressor variable depending on the type of performance predicted.

Contributions to Basic Science

This project advances our understanding of job performance prediction by providing evidence in support of a theory of the process by which individual-difference variables predict specific dimensions of performance. Many theories of the relationship between individual differences and job performance have been proposed, but there has been very little empirical support for any of them. This is especially true for less-studied aspects of performance such as citizenship performance and adaptive performance, which were the primary focus of this research. An established model of the process by which individual differences in predictor variables lead to individual differences in specific dimensions of performance can be used to guide research linking specific predictors to specific performance dimensions by helping to identify theoretically relevant predictors for different criteria. In the model tested, the construct of motivation was expanded to highlight how different predictors influence different components of motivation. All three components of motivation (motives, proactive cognitive processes, and self-regulation) were direct and/or indirect determinants of performance, depending on the performance dimension predicted. Each of these components, as well as knowledge and skill, was predicted by different individual difference variables, demonstrating the different paths through which individual differences influence performance.

This project also yielded several potentially useful products for the Army and future research. We created a multi-source performance rating instrument to measure task, citizenship, and adaptive performance. These performance components are important aspects of the Army junior commissioned officer performance construct domain. We created a situational judgment test (SJT) to assess knowledge relevant to the performance dimensions chosen for study. The SJT presents the individual with a situation and a set of response options varying in their effectiveness. We developed a past behavior record to measure skill relevant to the targeted performance dimensions. Using a highly structured format to minimize the possible confounding effect of writing ability, cadets described a recent situation they experienced that elicited behavior relevant to the performance dimension of interest and the behavior they exhibited. Trained raters used standardized rating scales to evaluate the level of skill that behavior represents. By allowing respondents to describe their best example of behavior relevant to a performance dimension, we can assess the extent to which the person is able to exhibit the behavior irrespective of the extent to which they typically exhibit the behavior. We demonstrated that this instrument has criterion-related validity, making it a quicker and less expensive alternative to a structured interview. To measure the components of motivation described in Johnson's (2003) model, we created (a) a unique self-report measure of work habits; (b) a motives scale assessing attitudes, values, and preferences that predict performance; (c) a cognitive processes motivation scale measuring self-efficacy, expectancy, goal content, and goal commitment; and (d) a self-report measure of self-regulation with respect to targeted behaviors.

Potential Army/Military Applications

This performance model provides the Army with a tool that can be used to identify the constructs through which individual difference variables work to influence performance on specific dimensions. This is especially important for citizenship and adaptive performance, which are important components of Army officer jobs but have received little research attention in a military or civilian context. This tool will contribute to a better understanding of the relationships between predictors, mediators, and job performance criteria. For the Army, this tool would probably have its best application in identifying training and/or development needs. Given a criterion construct on which an individual's performance is in need of improvement, this model can help to identify the determinants of performance on that construct. For example, an individual possessing adequate skill and knowledge may determine that he or she must learn new self-regulatory strategies to maintain goal-directed behavior.

Several products that may be very useful to the Army were developed for this project. First, the various motivation-related instruments we developed/assembled and validated against criteria of importance to the Army should prove to be quite useful as self-development tools. Completing these instruments can help Soldiers and/or officers learn about their own attitudes, values, goal commitment, self-efficacy, self-regulatory skills, and habits relevant to five critical performance dimensions.

Second, the Past Behavior Record Form may prove to be useful for training. In particular, the evaluation guide used to score this instrument will provide useful information regarding where various skill-relevant behaviors fall on various performance continua. Moreover, the principles induced during the process of calibrating the performance-relevant skills to different

points on these anchored rating scales will make useful learning points in training programs. The situational judgment test will also be a useful self-paced training tool. Individuals are presented with realistic situations and choose from a variety of alternative response options. Explanations of the strengths and weaknesses of each response can be developed to help the trainee understand the level of effectiveness of the chosen option and the most effective option.

Finally, the performance rating form may have applied value for performance appraisal, training needs analysis, or criteria for future validation studies. In addition, it would make a useful feedback and development tool. The performance rating form is a multi-source rating instrument. As such, it can not only provide a profile of strengths and weaknesses, it can provide information about discrepancies between how an individual is seen from different perspectives (i.e., subordinates, peers, superiors). It would be extremely useful, for example, for a junior officer to become aware of differences between his or her self-perceived performance and the perceptions of his or her performance held by subordinates.

Final Summary

Organizational citizenship performance and adaptive performance are increasingly important aspects of Army junior commissioned officer job performance. There has been very little research studying these variables in a military setting, and there is no well-established theory of the process by which individual differences predict them. In this research, we tested a theory of the process through which individual difference variables work to influence performance on specific citizenship and adaptive performance dimensions relevant to junior commissioned officers in the U. S. Army.

In order to test this theory, we assembled and developed a battery of instruments that are construct-valid measures of each component of the theory. The predictor battery was administered to 155 ROTC cadets in the Fall of 2006. Performance ratings were collected from one to five peers for each cadet during March and April of 2007, allowing us to assess the extent to which cadets performed the behaviors they intended to perform. A self-regulation scale was administered four times during this period, allowing us to track how well each cadet was able to maintain focus on performing these behaviors.

We found that modified versions of the proposed model described the process by which predictor variables influence performance. Different models were necessary to describe the process for task, citizenship, and adaptive performance. The strongest support for the model came from the task performance dimension. For this dimension, knowledge, skill, cognitive choice aspects of motivation, and self-regulation were all direct determinants of performance, and each of these determinants was predicted by different combinations of personality traits, experience, and motives. Motives such as job satisfaction and affective commitment were related to performance only through the mediating influence of expectancies and goal commitment. A similar model was supported for the adaptive performance dimension. The differences were that only some of the hypothesized direct determinants were significantly related to performance, and one personality trait had a direct influence on performance.

This model did not generalize to the citizenship performance dimensions, although a similar model was found for each different dimension of citizenship performance. In this model, motives were a direct determinant of performance and expectancies were not. Knowledge, skill, and self-regulation were direct determinants, but never all at once and the one that was a determinant depended on the citizenship dimension of interest. There were also two cases in which a hypothesized indirect performance determinant had a direct influence on performance.

There are several implications of these results. First, we have made great progress in identifying the mediators through which personality, ability, and experience operate to influence individual job performance. These individual differences lead to job performance because they influence one or more of the more proximal performance determinants of knowledge, skill, cognitive choice aspects of motivation, self-regulation, and motives. Second, there appear to be fundamental differences in how the model operates depending on the type of performance predicted. For example, the model that describes how individual differences influence task performance is different from the model that describes how individual differences influence citizenship performance. Motives such as job satisfaction directly influence citizenship performance but only indirectly influence task performance. There are also more direct determinants of task performance than of citizenship performance. Finally, the influence of motivation on performance is far more complex than suggested by previous models. There are several components of motivation and each component may be a direct determinant, an indirect determinant, a mediator, or a suppressor variable depending on the type of performance predicted.

Based on these results, we recommend that the Army use this performance model as a tool to identify the constructs through which individual difference variables work to influence performance on specific dimensions. An obvious application is in selection, because the model can help to identify predictors that will most successfully predict performance on whatever dimensions in which the Army is interested. Another application is in identifying training and/or development needs. Given a criterion construct on which an individual's performance is in need of improvement, this model can help to identify the determinants of performance on that construct. For example, an individual possessing adequate skill and knowledge may determine that he or she must learn new self-regulatory strategies to maintain goal-directed behavior.

We also recommend that the Army use the products that were developed for this project. First, the various motivation-related instruments we developed/assembled and validated against criteria of importance to the Army should prove to be quite useful as self-development tools. Completing these instruments can help Soldiers and/or officers learn about their own attitudes, values, goal commitment, self-efficacy, self-regulatory skills, and habits relevant to five critical performance dimensions. Second, the Past Behavior Record Form may prove to be useful for training. In particular, the evaluation guide used to score this instrument will provide useful information regarding where various skill-relevant behaviors fall on various performance continua. Moreover, the principles induced during the process of calibrating the performance-relevant skills to different points on these anchored rating scales will make useful learning points in training programs. The situational judgment test will also be a useful self-paced training tool. Individuals are presented with realistic situations and choose from a variety of alternative response options. Explanations of the strengths and weaknesses of each response can be

developed to help the trainee understand the level of effectiveness of the chosen option and the most effective option. Finally, the performance rating form may have applied value for performance appraisal, training needs analysis, or criteria for future validation studies. In addition, it would make a useful feedback and development tool. The performance rating form is a multi-source rating instrument. As such, it can not only provide a profile of strengths and weaknesses, it can provide information about discrepancies between how an individual is seen from different perspectives (i.e., subordinates, peers, superiors). It would be extremely useful, for example, for a junior officer to become aware of differences between his or her self-perceived performance and the perceptions of his or her performance held by subordinates.

There are a number of areas of future research suggested by this project, some of which can be conducted using the data already collected. One area of research we are planning is the extent to which different methods of scoring SJTs produce comparable scores, reliabilities, and criterion-related validities. The different methods that can be studied are (a) scoring based on a correct-incorrect coding of the "best" response only, (b) scoring based on the mean effectiveness rating of the response chosen, (c) scoring based on selection of both a "best" and a "worst" response, and (d) scoring based on rating the effectiveness level of each response. Each of these scoring procedures can be evaluated for (a) consensus-based scoring, (b) expert-based scoring, and (c) empirical scoring. Consensus-based scoring allows one to examine differences in effectiveness ratings associated with subgroups of examinees, so it is possible to determine if there are meaningful differences in the appropriate situational responses across different ROTC branches or across third- and fourth-year cadets/midshipmen.

One implication of the performance models we identified is that narrow personality traits should be better predictors of specific performance dimensions than should broad personality variables such as the Big Five personality traits. The determinants of performance are necessarily at a more specific level than is the performance dimension itself, and prediction is better when constructs are matched on level of specificity (Schneider, Hough, & Dunnette, 1996). The data collected in this research and the support for the models tested can be used to explain the predictive superiority of narrow personality traits by demonstrating that they have stronger relationships with the determinants of performance. For example, the anger facet of Emotional Stability was expected to influence more determinants of maintaining good working relationships than was the anxiety facet. This can be demonstrated by examining the correlations of each facet with the different variables that have been found to influence performance. Facets that are correlated with more determinants of performance should have larger correlations with performance, and when this happens the facet score also should be a better predictor than the factor score.

Another direction for future research is to study how potential moderator variables influence the relationships between performance and its determinants. Some of the moderators that have been found to influence the extent to which personality predicts performance are situational strength, occupation, time on job, autonomy, and typical versus maximum performance measurement. Personality and ability have been found to interact when predicting performance, although most recent investigations have shown no interaction.

We remain convinced that work habits can directly influence performance in some situations even though that was not found for the performance dimensions we studied. Work habits should influence performance directly when job-relevant behavior occurs automatically despite motivation to behave otherwise. Autonomy is expected to moderate this direct relationship because habits will have less of an influence on performance in stronger situations. Autonomy is the extent to which the environment allows an individual to behave in idiosyncratic ways. The stronger a situation, the less autonomy the individual has. The ROTC environment may be too strong a situation for habits to be related to performance, so future research could examine this relationship in an environment that clearly allows for the operation of habits. In addition, the criterion must consist of behaviors that can occur automatically. Given a criterion like this, the opportunity for habits to influence performance, and variance in the extent to which individuals' habits are consistent with performance, we expect that habits influence performance and self-regulation in the manner hypothesized.

Future research should explore alternative ways of measuring skill, especially for constructs like maintaining good working relationships and adaptive performance. Structured interviews, work sample tests, role plays, and computerized simulations are alternative ways of measuring skills that would probably provide better measurement than the Past Behavior Record.

This research should be replicated with different performance dimensions. The performance model was consistent across the three citizenship performance dimensions included in this project, so it would be beneficial to test the model on other citizenship performance dimensions. Another type of performance that was not addressed in this research is counterproductive work behavior (CWB). Models of the relationships between individual differences and CWBs are generally consistent with the citizenship performance model for which we found support, so extending this research to CWBs is a natural next step.

Finally, this research should be replicated in different samples. We used an ROTC sample because they are in training to become junior commissioned officers and we expected the results to generalize. Future research should be conducted in a sample of actual junior commissioned officers to ensure these results are consistent across samples. Enlisted Soldiers, NCOs, and senior officers are other potential military samples, and the model also can be tested in civilian samples for different types of jobs.

ARI has created various temperament scales thought to be related to adaptive performance and has empirically linked these measures to the mission performance of Special Forces (SF) Soldiers. ARI is now leveraging this measure to create a self-development feedback packet for SF Officers. The performance model identified for predicting adaptive performance in this research can be used to evaluate the temperament scales currently included and to identify additional or alternative measures for this feedback packet. The measures created in this project are designed to predict performance at a more proximal level than the current temperament scales. The Situational Judgment Test (SJT) and Past Behavior Record (PBR) Evaluation Guide developed as part of this project would be ideal additions to this type of package. With the SJT, individuals are presented with realistic situations and choose from a variety of alternative response options. After completing each SJT item, a Soldier could be provided with the rated level of effectiveness for each response option, along with an explanation for the strengths and

weaknesses of each response. The PBR evaluation guide can provide useful information regarding where various skill-relevant behaviors fall on various performance continuums. Moreover, the principles induced during the process of calibrating the performance-relevant skills to different points on these anchored rating scales will make useful learning points.

This model and the measures developed in this project could be used to create, revise, or improve the SF self-development feedback packages or other Soldier or officer self-development packages that may be used or developed outside of SF. For example, self-development packages could be created to improve any type of citizenship performance in addition to adaptive performance. The model and results of this basic research project can be used to identify relevant predictors for five different performance dimensions (decision making/problem solving, maintaining good working relationships, organizational commitment, showing initiative, and adapting to uncertain or changing situations), and the products developed for this project can be incorporated into self-development feedback packages.

References

Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 35-70). San Francisco: Jossey-Bass.

Johnson, J. W. (2003). Toward a better understanding of the relationship between personality and individual job performance. In M. R. Barrick & A. M. Ryan (Eds.), *Personality and work: Reconsidering the role of personality in organizations* (pp. 83-120). New York: Jossey-Bass.

Motowidlo, S. J., Borman, W. C., & Schmit, M. J. (1997). A theory of individual differences in task and contextual performance. *Human Performance, 10,* 71-83.

Schneider, R. J., Hough, L. M., & Dunnette, M. D. (1996). Broadsided by broad traits: How to sink science in five dimensions or less. *Journal of Organizational Behavior*, *17*, 639-655.

Goal-Driven Perception and Cognition in Complex Social Environments

Contract #: W74V8H-05-K-0003
Institution: Arizona State University

Contract Dates: 01/01/2005 to 12/31/2007
PI: Douglas T. Kenrick
Co-PI: Steven L. Neuberg

Problem(s)/Research Question(s) – How do common social goals relevant to military situations (e.g., self-protection, other-protection, retribution) influence early-stage cognitive processing of complex situations? To which person in a crowd will we pay attention, for example, and how will we interpret his or her facial expression? How do these goal-driven processes shape important down-stream behavioral decisions (e.g., to shoot vs. not to shoot a possible enemy combatant)?

Technical Barrier(s) – Until recently, little effort has been made to link theory about fundamental social goals and their associated emotions to theory about how people process information during the first seconds of an initial encounter.

Significance/Impact for Basic Research – This research will help to conceptually integrate two important, but disconnected, research literatures while also providing empirical data that tests hypotheses about goal influences on cognitive processing of complex social environments.

Potential Transitions – Knowledge gained in this basic research may be useful in identifying individuals susceptible or resistant to particular goal-related processing biases, and designing training programs to mitigate such biases. Our research may be most directly relevant to the following applied programs:

- The Army's Science and Technology Objective: Soldier (STO IV.SP.2002.01)—Selection, Classification, and Performance Metrics for the Future Force Soldier,
- The Army's Science and Technology Objective: Training (STO IV.SP.2003.06)—Objective Force Warrior Training,
- The Army's Science and Technology Objective: Training (extension of Future Force Warrior Training)—Maximizing Small Unit Performance,
- The Army's Science and Technology Objective: Training—Ground Systems Team Training,
- The Army's Science and Technology Objective: Training (STO IV.SP.2003.06)—VICTOR: Virtual Individual and Collective Training for Future Warriors, and
- The Army's Science and Technology Objective: Leadership Development—Assessing and Developing Leaders for Special Operations Units.

Overview

Soldiers often encounter extremely complex social situations in which they must rapidly make difficult, life-or-death decisions. However, human cognitive capacities are limited in ways that make it impossible to attend equally to every individual or even to all characteristics of any given individual encountered in such a setting (imagine a crowded market in a large city).

Instead, people must focus their attention and cognitive resources on a small subset of individuals and characteristics. The selective direction of cognitive processing often occurs automatically, without conscious intent, and is highly susceptible to biases that can potentially lead to serious errors (e.g., failing to process clues of an impending fatal ambush or mistaking an innocent civilian for an armed insurgent). Furthermore, a person's current motivational state may change the ways in which that person allocates attention and other limited cognitive resources. The aim of this research program is to assess how common social goals—self-protection, other-protection, and retribution—might shape attention and information processing in complex social environments and thereby influence important down-stream behavioral decisions.

Research Approach

We proposed and ran 10 laboratory experiments to test hypotheses derived from our conceptual framework. Each experiment used established research methods from social psychology and cognitive science designed to tap different levels of processing—from visual attention (measured via eye-tracking) and interpretation through memory and decision-making (e.g., tasks requiring rapid friend/enemy distinctions and shoot/don't-shoot decisions). In addition, each effort included manipulations of participants' goals (self-protection, other-protection, and/or retribution). Besides experimental manipulations, we assessed chronic individual differences in these goals. Participants were confronted with complex social scenes in which target individuals differ in their gender, age (combatant age or not), group membership (ingroup, outgroup, or outgroup ally), facial expression (angry or not), and/or possession of a weapon. Our ongoing analyses are examining how these different goals alter visual attention to, encoding of, memory for, and/or behavioral decisions about the different individuals in the scenes.

Accomplishments

Thus far, we have completed data collection for all 10 proposed experiments, and are currently completing data analysis, integrating findings across efforts, and preparing research papers describing the interesting results. To overview, results indicate distinctive patterns of cognitive bias in the processing of adult Arab males, as well as several interesting interactions between cognitive processing and participants' motivational states. Participants in control conditions are generally not good at distinguishing adult Arab males they have seen from those they have not seen. Furthermore, adult Arab males are most likely to be labeled as enemies (even when they are not), and to have anger attributed to them (even when they are displaying neutral facial expressions). When any one of various forms of threat are aroused, people become better at distinguishing among individual Arab males. Men who are experiencing revenge motivation, although also more attentive to individual Arab males, are also quicker to decide to shoot those individuals in a rapid decision-making task.

Two of our experiments (Experiments 1 and 2) were designed to assess how social motivations (i.e., feeling self-protective, other-protective, or vengeful) would influence visual attention measured via eye-tracker. We recorded how participants visually scanned a group of people (including men, women, boys and girls, who were Arab or European, and who were wearing angry or neutral facial expressions). We later asked participants to identify particular

faces that had been in those arrays, and the test included an equal number of people in each category that they had not seen before. This combination of tasks allows us to examine processing efficiency, or "bang for the attentional buck," by calculating the relationship between amount of time spent looking at a particular type of face and later memory for faces in that category. One interesting finding from that research is that women concerned with self-protection become increasingly efficient at processing Arab males.

This year we replicated this finding of enhanced processing efficiency with a different outgroup for our European American perceivers: African American males. We found both genders to be more efficient at processing African American males after a self-protection manipulation.

What mechanisms underlie the increased processing efficiency of Arab males? One possibility is that Arab males are *encoded* more efficiently than are other face types—that they are more efficiently set into memory. Alternatively, encoding may be no more efficient for Arab males than for other face types, but memory for their faces may be *more resistant to decay* once stored in long-term memory. We are currently planning an effort to differentiate between these alternatives by investigating short-term rather than long-term memory processes. Specifically, if Arab males are more rapidly retrieved from short-term memory than comparison target individuals, this would suggest that they are encoded more efficiently than the other face types. If, instead, Arab males are no more rapidly retrieved from short-term memory than other face types, this would suggest they are more resistant to decay in long-term memory.

Experiments 3 and 4 were designed to assess how the same social motivations affect rapid categorization of people as enemies or allies. Each participant viewed 128 photos varying by gender, group (European-American vs. apparently Arab), age (mid-20s vs. elementary-age children), and facial expression (angry vs. neutral). Half the targets were randomly labeled with an ingroup insignia; the others were randomly labeled with the enemy insignia. Data analysis for Experiment 3 suggests that participants were generally biased toward identifying Arabs as enemies and Europeans as friends; they also were biased toward identifying females as friends. When concerned with self-protection, the bias to identify Arabs as enemies increased. Moreover, participants were slower to react to angry Arab faces, and were especially biased to call them enemies even when they were not; this was particularly the case for participants with implicit prejudices against Arabs. In contrast, participants were quick to react to neutral European faces, and were especially biased to call them friends even when they were not; this was particularly the case for participants who are inclined to see the world as a dangerous place. Revenge motivations also seem to create a bias to identify ingroup members as friends (even when they were not); this bias was pronounced among participants who dispositionally view the world as a dangerous place. A manuscript based on these results is being prepared.

In a follow-up effort, we exposed participants to Arab and European male and female faces that were either all angry or all neutral. In the past, we observed strong effects of emotional expressions (i.e., anger) on judgments of faces as enemies or friends. By exposing participants to only one type of facial expression, we can detect more subtle effects that may have been previously obscured by the mixture of several emotional expressions. Data from two experiments now converge on the conclusion that participants are more biased to call Arab men enemies than

they were to call Arab women enemies. Conversely, they are biased to call European-American women friends.

Experiments 5 and 6 were designed to assess how social motivations affect attributions of anger to neutrally-expressive faces of different groups. The most interesting finding from this research is that attibutions of anger are high for targets who are Arab, who are men, and who are adult, and that these features contribute to anger attributions in a straightforward additive fashion. That is, judges tended to attribute higher levels of anger to adult Arab males (who possess the three cues to danger) than to, for example, adult Arab females or adult Euro males (who possess two cues to danger)—even when all targets were, in fact, displaying neutral facial expressions.

Experiments 7 and 8 were designed to assess how the presence of individuals relevant to some motives might interfere with the cognitive processing of individuals relevant to other goals. In these experiments, participants viewed pairs of photographs including either an adult Arab male, an adult European-origin female, or a crying child. The goal was to examine (1) whether and in what way these stimuli interfere with processing other faces presented alongside them, and (2) whether the nature of this interference differs depending on the perceiver's active and chronic goals. Whereas participants had an especially difficult time accurately differentiating Arab targets from one another in memory under baseline conditions, each of the threat manipulations improved this performance. Morever, each of the threat motivations created a tendency for Arab male targets to "trump" the recognition of white female targets (i.e., to make it more difficult for participants to differentiate in memory European-origin females from one another); in the retribution condition, Arab male targets also "trumped" recognition memory for (the normally attention-grabbing) crying children. For our participants, the presence of Arab males appeared to "steal" processing away from other target individuals in the social context.

Experiments 9 and 10 were designed to assess how the three motivations affect rapid categorization of other people as threats or non-threats, and the decision to shoot those carrying a weapon. Each participant viewed photos varying by gender, group (European-American vs. apparently Arab), and facial expression (angry vs. neutral). Half the targets were holding a weapon, and the other half were holding a wallet or a cell phone. The participant's task was to shoot only armed targets, regardless of gender, ethnicity, or emotion. When motivated by retribution, participants were especially likely to not shoot *unarmed* Europeans and to shoot *armed* Arabs, but they also exhibited a bias to shoot unarmed Arabs. These findings are scheduled to be presented this winter at the annual meeting of the Society for Personality and Social Psychology.

Contributions to Basic Research

This research promises to make several important contributions: (1) It theoretically links ecologically inspired theories of motivation and cognition with research on fundamental human goals and information processing; (2) Findings can contribute to a theoretical framework for understanding how goals related to self-protection, other-protection, and retribution shape early-stage attention and interpretation of complex, information rich social environments, and thereby influence subsequent behavioral decision making; (3) We have been able to provide rigorous

empirical tests of the hypotheses derived from this framework; and (4) The findings have suggested a number of interesting new questions worth pursuing. Thus, this research should help expand upon the field's understanding of motivated social perception and cognition.

Potential Army/Military Applications

Several of these findings are relevant to understanding how a Soldier's perceptual, cognitive, and decision making processes may depend, for example, on whether he or she is currently concerned with self-protection, protection of a friend or comrade, retribution for an attack on one's country or the death a comrade, or the safety and well-being of innocents in the field. Although not all data are yet analyzed, initial results do have direct relevance to Soldiers making judgments in conflict situations. The findings that judges are generally not good at distinguishing Arab males from one another, are disposed to label those individuals as angry and as enemies even when they are not, and are quicker to make a decision to shoot those individuals are of direct relevance to Soldiers involved in Middle Eastern conflicts who must in fact distinguish between friendly and hostile Arabs. Given that combatants may include both men and women, it is also of interest that revenge motivation increased the proclivity of men to shoot Arab males, whereas self-protection increased the proclivity of women to process Arab males efficiently and accurately.

Future Plans

Several preliminary findings from this program have been presented at national conferences and are currently being developed into manuscripts for publication. We also are conducting additional analyses of the sometimes complex data sets, such as linking eye-tracking and memory findings, as well as exploring individual differences. We expect these more thorough analyses to uncover additional interesting patterns. Several potential follow-up efforts are suggested by the current patterns of findings. Possible lines of follow-up research involve these questions: (1) Are there individual differences in the patterns of cognitive biases such that individuals may be characterized by "syndromes" of biases, perhaps related to chronic sensitivities to fundamental motives, and might such biases influence how these individuals process information in complex social environments? Such a question would seem to have important implications for how one might identify military personnel susceptible to particular decision errors (or high quality decisions) in the field. One approach to exploring this would be to administer batteries of all our different measures and tasks to the same subjects across multiple measurement sessions, thereby enabling us to identify social-cognitive processing "types." (2) Given that there are sets of biases linked to particular motivational states, how would one train them away? Would the training approach differ depending on the particular patterns (syndromes) of bias individuals exhibited? (3) The types of cognitive biases we have demonstrated here will likely also be exhibited by those in contact with U.S. Soldiers, such as civilians and enemy combatants in foreign lands. Given what we are learning about how the biases are elicited, and their motivational underpinnings, what strategies might be useful to lessen the likelihood, or reduce the impact these biases have on the perceptions and treatment of these Soldiers? (4) How are personal vulnerabilities to specific threats affected by experiences in a war zone? Although personality is typically considered stable over time, we have found that some threat-focused personality measures, such as participants' belief that the world is a

dangerous place, change after participants are exposed to our motivation manipulations (e.g., self-protection threat). What implications does this have for Soldiers returning home from a war zone? What distinguishes those individuals who become habituated to the constant state of arousal in the war zone from those who become increasingly sensitive to threat?

An Integration of Motivation Theories

Contract #: DASW01-04-K-0001
Institution: The Hebrew University of Jerusalem

Contract Dates: 12/12/2003 to 12/11/2007
PI: Avraham N. Kluger
Liaisons: Trueman Tremble, SARU
Robert Solick, LDRU

Problem(s)/Research Question(s) – Can "amity goals" protect people's performance from the negative psychological consequences of failure?

Technical Barrier(s) – Performing multiple experiments to test each step of the proposal with relatively large scale samples.

Significance/Impact for Basic Research – Showing how a single concept of "amity goals" changes our understanding of performance antecedents.

Potential Transitions – Knowledge gained during our 2007 work may be usefully employed in:

- Various Army training programs,
- Performance enhancing, and
- LEADERSHIP: Developing leaders in a changing Army (WP 103).

Overview

Some people believe intelligence is fixed while others believe it is malleable. Research in Social Psychology has shown that people's beliefs about intelligence affect how they construe various achievement situations and hence their performance, especially following failure. Those who believe that intelligence is a fixed entity view achievement situations as tests of their ability. They develop "performance goals"—the goal to perform well that proves that one is not stupid. In contrast, those who believe that intelligence is malleable view achievement situations as opportunities to develop one's skills. They develop "mastery goals"—the goal to increase one's intelligence through experience. Thus, it is not surprising that when faced with failure, those who developed a performance goal feel stupid; they bolster their belief that they lack ability, they reduce their effort, and hence their performance deteriorates. In contrast, those who developed mastery goals see a failure as an opportunity to learn; their effort remains intact or increases and hence their performance may be improved. This is the essence of goal orientation theory (Dweck & Legget, 1988), which optimistically shows that although people naturally differ from each other with respect to their adoption of mastery goals, adoption of mastery goals can be induced with instruction. However, recent theory development describes three rather than two goals (e.g., Vandewalle, 1997; Elliot 1999): mastery goals (the goal to develop competence), performance-approach (or proving) goals (the goal to demonstrate competence), and performance-avoidance goals (the goal to avoid the demonstration of lack of competence).

In an effort to integrate motivation theories, we propose that people tend toward one of these three goals partly to address their basic human values (Schwartz, 1992). That is, we are

making a link between a theory that is specific to achievement situations (goal-orientation theory) with a generic motivation theory regarding human values (value theory).

Values are considered to be desirable, trans-situational goals reflecting both biological needs and conformity to social pressures that serve as guiding principles in one's life (Schwartz, Lehmann, & Roccas, 1999). Schwartz's Values Framework postulates 10 values types organized as a circle: security, conformity, tradition, benevolence, universalism, self-direction, stimulation, hedonism, achievement, and power. Adjacent values tend to be compatible whereas values in opposing positions tend to be in conflict. For example, "the pursuit of novelty and change (stimulation) is likely to undermine preservation of time-honored customs (tradition)" (Schwartz, Lehmann, & Roccas, 1999, p. 110). The values circle is organized by two sets of higher-order values that are in conflict: change (stimulation and self-direction) versus conservation (security, conformity, and tradition) and self-enhancement (power, achievement, and hedonism) versus self-transcendence (universalism and benevolence).

Considering these two theories, we suggest that mastery goals are related to stimulation and self direction values; performance-avoidance goals are related to security, conformity, and tradition; and values and performance-approach goals are related to power and achievement values. Furthermore, through this comparison, we identified a family of values that are not reflected in current goal orientation research and that hint at the existence of a fourth goal orientation: amity goals. We suggest that self transcendence values (e.g., benevolence) may serve as antecedents of amity goals. Furthermore, we suggest that mastery goals and performance-avoidance goals' main concern is the fulfillment of the need for competence while performance-approach goals (or contest goals) and amity goals' main concern is the fulfillment of the need for relatedness. This structure reflects the structure of values with its two axes of (1) openness to change versus conservation and (2) self-enhancement versus self transcendence. (Note that competence and relatedness goals are the foundation of many other motivational theories.)

One important benefit of the four goal orientations model is that it clarifies the complex reactions to negative feedback. We suggest that the juxtaposition of amity goals with performance-avoidance goals immunize those who have performance-avoidance goals against its detrimental effects on emotions, motivation and performance, and that the juxtaposition of amity goals with mastery goals creates the optimal setting for achievement.

In summary, we propose that the amity goals (even in competitive environments) are (1) missing from leading performance theory, (2) provide an emotional buffer against failure, and (3) shield the performance of those who adopt them from detrimental effects of failure through the two mechanisms of valence and expectancy.

Research Approach

To collect evidence for the existence of amity goals, we used three questionnaires that were analyzed with smallest space analysis (SSA), a variant of multiple dimensional scaling:

- A survey containing all the existing goal-orientation questionnaires and Schwartz's Short Values Survey (736 university students),

- A survey containing all the above with new amity-goals items (145 business school undergraduates), and
- A survey containing all the existing goal-orientation questionnaires with amity goals and life-basic needs (97 business school undergraduates).

To demonstrate the utility of the amity goals construct, we performed both a scenario experiment and an actual performance experiment:

- The experiment (187 university students) included four scenarios. In each scenario, two of the four goals were salient. Here are two of the four scenarios as examples: *Mastery and amity scenario*: "Imagine you enrolled to a challenging class that enables you to learn a lot. You desire to completely master the material presented in this class. It is important for you to assist your friends to succeed in this class. You have failed this class's midterm exam. This exam counts for 40% of the final grade." *Performance-avoidance and contest scenario*: "Imagine you enrolled in a class you feel you have a good chance to succeed in. Your goal for this class is to avoid performing poorly compared to the rest of the class. It is important for you to do better than other students in this class. You have failed this class's mid term exam. This exam counts for 40% of the final grade." Motivation following the failure was measured using two items (Vroom, 1964), one that measures expectancy ("In comparison to the expectations you had to succeed in this course while enrolling in it, what are your expectations now?") and one that measures valence ("In comparison to the value this course had for you while enrolling in it, what is its value now?"), on a scale from -5 (much lower) to 5 (much higher).

- Undergraduate students (N=133) attending small classes of approximately 10 students each were randomly assigned in a 2 (contest vs. amity) x 2 (mastery vs. performance avoidance) experiment. For example, for the amity manipulation, students were asked to choose a partner and tell the partner two surprising things about themselves and to invent together a lie about each other. Next they had to tell the group three things about each person, where one was false and the other two were true. The group had to guess the lie. Following the manipulation, each participant was asked to invent as many uses for a broomstick as possible in 5 minutes. Next they got a filler task, followed by bogus feedback ("most students found X usages"). For 110 participants, the "X" was above their own performance (only their data is reported here). Next, they were asked to rate the feedback and both their expectancy to succeed in a similar task, and the importance of succeeding in a similar task. Finally, they were asked to invent as many uses for a wool bulb as possible, and then they were debriefed. Within a week they were asked to report how they felt regarding the whole experience.

Accomplishments

All efforts corroborated our proposal regarding amity goals. Here, due to space limitations, we present only a sample of our results. First, below is a map (Figure 1) obtained in our second questionnaire showing a spatial representation of the correlations among all goal-orientation items**.**
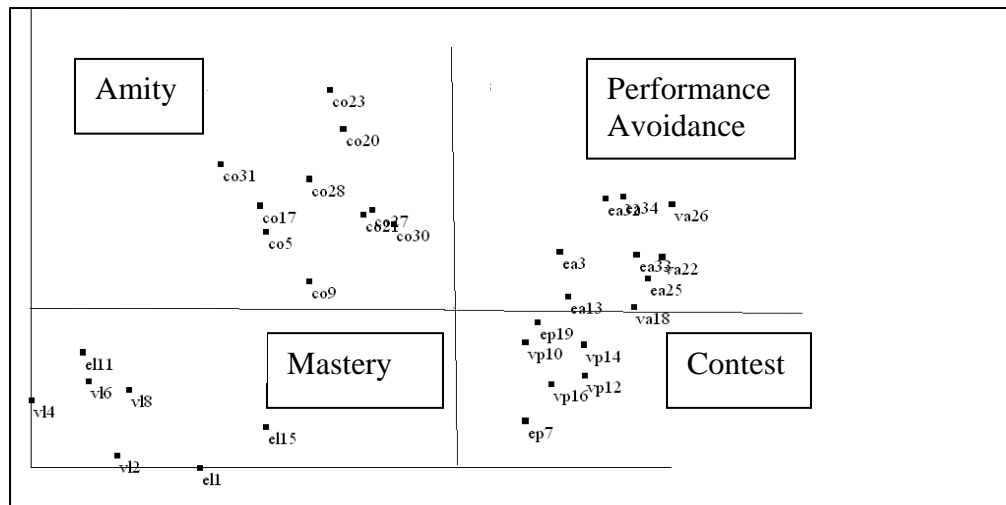
Figure 1. Smallest space analysis of items putatively measuring four goal orientations.

All the items marked with "co" are the new amity goal items. It is clear from this SSA map that items measuring amity goals are relatively homogenous and separable from all the other items used in the literature to measure goal orientations.

Figure 2 shows the correlations between the four different goal orientations and the 10 values, taken from the same research effort.
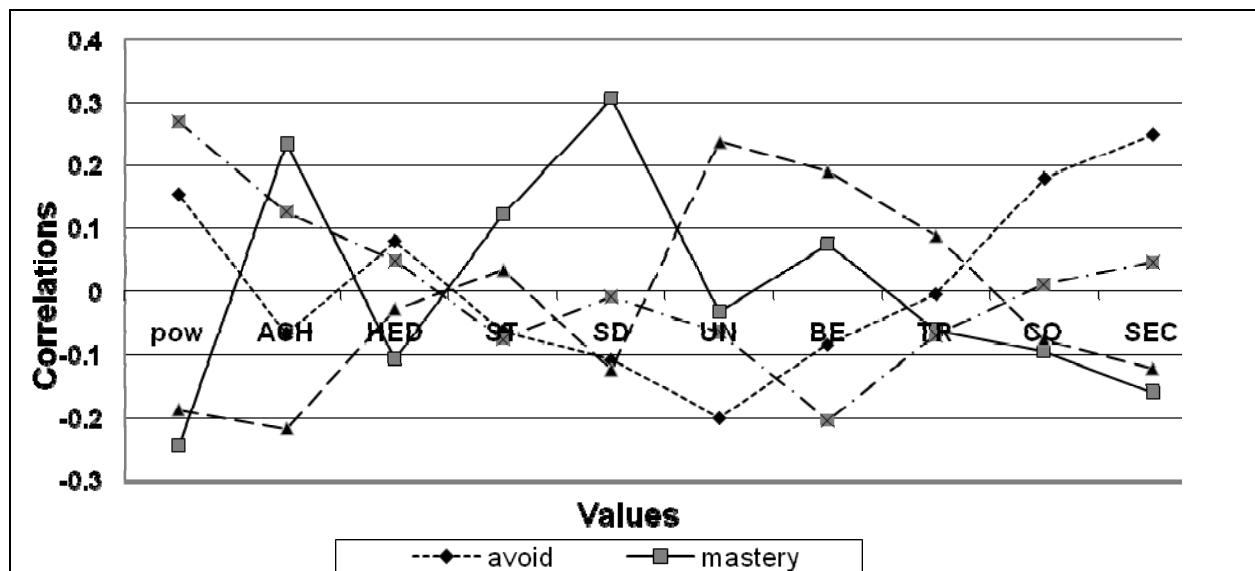

Figure 2. Correlations of values and goal orientations.

Note. pow=power, ACH=achievement, HED=hedonism, ST=stimulation, SD=self direction, UN=universalism, BE=benevolence, CO=conformity, TR=tradition, SEC=security.

As can be seen in Figure 2, each goal orientation shows highest positive correlations with the expected set of values. For example, amity goals are most positively correlated with universalism and benevolence, and most negatively correlated with achievement and power values. This and other set of results, not reported here, corroborate the validity and the meaning of our proposed construct of amity goals.
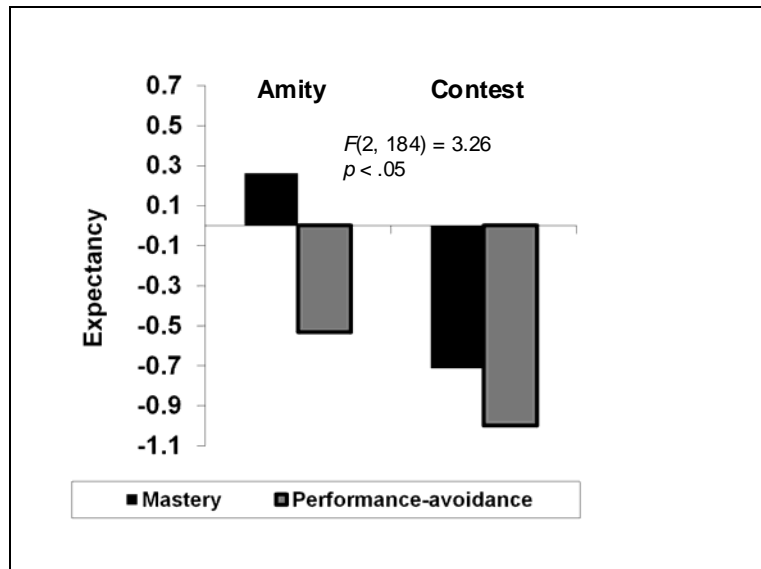
Figure 3. Expectancy to succeed following negative feedback.

Next, a manipulation of amity goals in a scenario experiment showed that combining amity goals with mastery goals produced the best possible response to negative feedback both in terms of expectancy to succeed (see Figure 3) and the value of succeeding.

Finally, the actual performance experiment showed that amity goals protect both the expectancy (see Figure 4) and the valence of participants who were induced to consider performance avoidance goals.

Figure 4. Expectancy to succeed following negative feedback on actual performance.

Contributions to Basic Research

This work contributes to basic research by introducing a new construct that adds to a growing body of knowledge regarding factors affecting performance in achievement situations, and by demonstrating its utility to produce conditions that both pro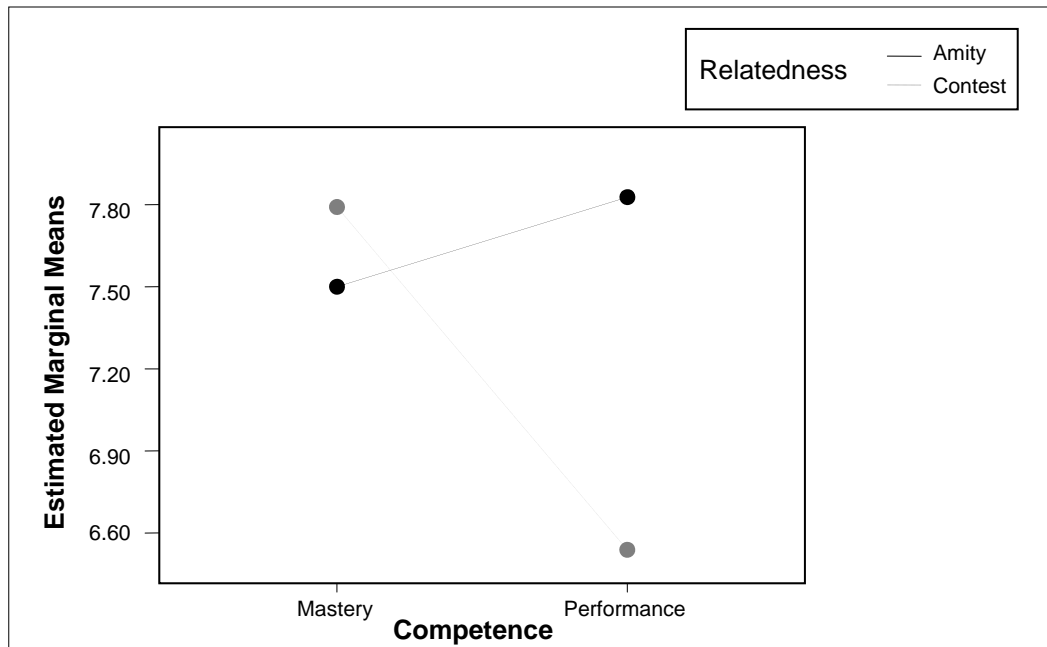tect one from the detrimental effects of negative feedback and augment the beneficial effects of adopting mastery goals.

Potential Army/Military Applications

We believe that training leaders to design settings with amity goals can be applied in various military settings. Below are three possible ways to apply amity goals in military settings. Given the early stage of this construct development, each of the following suggestions, if applied, needs to be monitored with field research.

- Increase amity goals in training sessions. To do so, one would have to probe for true combat stories during which amity goals, cooperation, and unit cohesiveness were contributing factors to operation success. These stories can be introduced into the training curricula to induce situational sensitivity to amity goals. Doing so may increase the learning achievements of the trainees.

- Use measures of amity goals in selection procedures. Measuring amity goals may be used to guarantee that each unit is populated with individuals with high amity goals to prevent the unit from becoming overly competitive (destructive for the learning capacity of the unit).

96

- Use knowledge of the benefit of amity goals in training leaders. For example, leaders can be encouraged to discover ways to increase amity goals in a unity. Possible means may include evaluating direct report on their mentoring of their peers and emphasizing the importance of helping one's peers with their learning.

Future Plans

Before the end of this research program, we plan to perform one more actual performance experiment to probe when amity goals protect against the debilitating effect of performance avoidance goals and when it augments the benefit of mastery goal. In addition, we plan to bring to completion three other dissertations not reported here: on feedforward (reported in 2005), on negotiational self (reported in 2006), and on attachment style and commitment (not reported yet).

<div align="center">

References

</div>

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256-273.

Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist, 34*(3), 169-189.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 25, pp. 1-65). New York: Academic Press.

Schwartz, S. H., Lehmann, A., & Roccas, S. (1999). Multimethod probes of basic human values. In J. Adamopoulos & Y. Kashima (Eds.), *Social Psychology and Culture Context: Essays in Honor of Harry C. Triandis* (pp. 107-123). Newbury Park, CA: Sage.

VandeWalle, D. (1997). Development and validation of a work domain goal orientation instrument. *Educational and Psychological Measurement, 57*(6), 995-1015.

Vroom, V. H. (1964). *Work and motivation*. New York: Wiley.

Exploring Pattern Recognition as a Predictor of Mental Flexibility

Contract #: W91WAW-07-C-0028 **Contract Dates:** 06/01/2007 to 05/31/2008
Institution: Wesleyan University **PI:** Cynthia Matthew
  Co-PI: Steven Stemler

Problem(s)/Research Question(s) – Does the basic ability to connect cues into meaningful patterns (pattern recognition) bring about flexible thinking in real-world situations?

Technical Barrier(s) – An accepted theory of pattern recognition that articulates its role and function in intelligent behavior does not exist. The literature suggests that pattern recognition may figure prominently in learning, judgment, and the capacity to shift "habits of mind"; however, there does not yet appear to be any research that explores how it relates to flexible thinking or mental flexibility.

Significance/Impact for Basic Research – Understanding the relationship between pattern recognition and mental flexibility may help explain why some people think more flexibly than others. This knowledge may ultimately lead to more effective assessment tools and training programs.

Potential Transitions – Applied research areas may include:

- Developing Leaders in a Changing Army, ARI Leadership Development Unit,
- Assessing and Developing Leaders for Special Operations Units,
- Advanced Technologies for Leader Development, and
- Technique and Tools for C4ISR Training of Future Brigade Combat Team Commanders and Staffs (Part of STO IV. SP. 2002. 02).

Overview

In the current military environment, problems and situations can differ fundamentally from those experienced in the past, and lessons learned from past experience and training may not readily apply in the world of today. Military officers must possess and develop the capacity to respond to novel types of problems and situations in new ways in real time. Traditional methods for assessing flexible thinking do not sufficiently explain why some people are more mentally flexible than others, even when they possess similar levels of mental ability. In addition, officers and Soldiers are bombarded with information. While in the past, the challenge was accessing information, the challenge of the future lies in the capacity to filter out unnecessary information, detect significant and meaningful patterns within the context of information rich environments, and successfully modify perceived patterns as information in the environment changes—all of which involve pattern recognition. The purpose of this research is to examine the relationship between pattern recognition and flexible thinking.

Pattern recognition is a basic building block of human intelligence. It refers to the thinking process of connecting cues or bits of relevant information to form meaningful patterns in a given situation. Depending on the situation, relevant cues may be, for example, elements in

the physical environment, pieces of technical information, or aspects of social behavior. A pattern recognition approach to problem solving and diagnosis has been identified as very successful when used by medical experts (Gilhooly, 1990; Coderre, Mandin, Harasym, & Fick, 2003). This effort investigates how pattern recognition may relate to mental flexibility at a broader, more basic level.

This research extends the work of a previous ARI-funded project, in which tests were developed to assess mental flexibility where findings showed that it seemed to be related to pattern recognition (DASWO-03-K-0001). For this research, we are utilizing some of the previously developed tests and have developed new tests and measures in the realm of social judgment and social influence. We expect to find that pattern recognition will bear a strong association with mental flexibility across social and non-social types of assessment instruments.

Research Approach

We are conducting a large-scale data collection from 200-300 college students recruited from three universities in the Northeast. The students will participate in three hours of testing that will include various surveys and tests that measure pattern recognition and mental flexibility, as well as traditional measures of intelligence. Using a sub-sample of participants, we also will obtain course performance ratings by their professors. Data obtained from these various measures will be statistically analyzed to test hypotheses that will evaluate pattern recognition as a basic ability that gives rise to flexible thinking, over and above traditional measures of intelligence.

Accomplishments

We are currently in the final stage of data collection and engaged in data entry and scoring. Our accomplishments to date include:

Test Development
- Developed and piloted three new tests of pattern recognition/mental flexibility:
  - Collegiate Social Judgment Test (assesses both Pattern Recognition and Mental Flexibility),
  - Social Influence Test (Pattern Recognition), and
  - Word Recognition Test (Pattern Recognition).
- Developed two new survey instruments:
  - Flexible Performance Survey (Self-rater and Faculty-rater versions), and
  - Background Survey (manifestations of real-world flexible performance).
- Conducted an advanced statistical analysis on two previously developed mental flexibility tests from an ARI project (#DASWO-03-K-0001) and reduced the number of questions on the Flexible Mapping and Counterfactual Analogies Tests.

Project Management
- Hired four part-time Research Assistants.
- Recruited at four local universities: Sacred Heart, Quinnipiac, Yale, and Central Connecticut State.

- Prepared and submitted Human Subjects Committee applications to Wesleyan and participating universities.
- Participated in interviews for media coverage of the project: Wesleyan Newsletter and CT NPR radio station.
- Set up an e-mail account and Web space for project administration and recruitment of participants.
- Developed all recruitment materials.

Participant Recruitment and Scheduling
- Sent e-mail announcements and scheduled volunteer participants at Sacred Heart (general and ROTC program) and Quinnipiac.
- Recruited classrooms at Wesleyan through direct contact with faculty (ongoing).
- Recruited participants through classroom presentations at Wesleyan (ongoing).
- Recruited classrooms through faculty e-mail at Central Connecticut State (project discontinued due to difficulties obtaining permission to use an on-campus facility).

Data Collection
- Completed administration for 213 participants from Sacred Heart and Quinnipiac Universities.
- Currently engaged in data collection at Wesleyan University.

Collaboration
- Participated in a telephone meeting with Dr. Robert Wisher, Director, Advanced Distributed Learning Office of the Secretary of Defense DUSD/Readiness and Training, at the Pentagon and Dr. Jan Cannon-Bowers at the University of Florida. The purpose of the meeting was to discuss collaborating on ways to bridge our research with applied training methods.
- Participated in a special conference on creativity and leadership held by the Society for Industrial and Organizational Psychology (October 24-26, 2008).

Contributions to Basic Science

This research has the potential to advance our understanding of what makes up mental flexibility/creative thinking, which will allow us to build more useful performance models and develop more accurate ways to assess it. In addition, we have developed new ways of testing pattern recognition and mental flexibility, specifically in the area of social relations, which will be available for future research. Finally, we have completed secondary analysis of two previously developed mental flexibility tests and have created a reduced set of questions for each that can be used for future research and applied purposes.

Potential Army/Military Applications

The results from this effort have implications for the training and development of future Army leaders. The implications may include developing military-specific simulations for the purpose of assessing and developing pattern recognition skills. In addition, adaptation of

methods used to develop the social judgment and social influence tests may be adapted to create instruments for training in military problem-solving.

Future Plans

Next year we plan to analyze and interpret our findings that will be summarized in a technical report, including an analysis of any implications and recommendations derived from this research. Should the results of our data analysis prove promising, we may propose extending the investigation for a second/third year to conduct experiments in the lab to assess the potential for applications to training.

References

Coderre, S., Mandin, H., Harasym, P. H., & Fick, G. H. (2003). Diagnostic reasoning strategies and diagnostic success. *Medical Education, 37*, 695-703.

Gilhooly, K. J. (1990). Cognitive psychology and medical diagnosis. *Applied Cognitive Psychology, 4*(4), 261-272.

The Development and Validation of a Situational Judgment Test to Predict Attrition Incrementally Over General Cognitive Ability and a Forced-Choice Personality Test

Contract #: W91WAW-07-C-0014   **Contract Dates:** 03/26/2007 to 03/25/2008
Institution: Work Skills First, Inc.   **PI:** Michael A. McDaniel

Problem(s)/Research Question(s) – Can personality and temperament be measured through a knowledge test? Can such a knowledge test be built so that respondents cannot improve their scores through faking? Will scores on the knowledge test predict attrition-related variables over and above cognitive ability and a personality test?

Technical Barrier(s) – Most attempts to measure personality and temperament in a faking-resistant manner have failed.

Significance/Impact for Basic Research – Understanding how to measure personality and temperament in a faking-resistant manner will lead to better prediction of human performance in work settings.

Potential Transitions – Knowledge gained in this research may be used in:

- ARI's applied training research programs,
- ARI's applied selection research programs, and
- Efforts to obtain a more demographically diverse pool of recruits.

Overview

People lie. People often lie when an important outcome can be achieved by lying, when lying is easy to do, and when one cannot be caught lying. This is the dilemma facing psychologists seeking to screen applicant pools on personality and temperament traits, such as adaptability, conscientiousness, agreeableness, and emotional stability. These traits should predict job and training performance as well as attrition. However, to collect such data on an applicant, one must ask the applicant for a self-report of their personality. Questions such as "Are you dependable?" would be excellent predictors of dependability if the applicants responded truthfully. However, many applicants do not respond truthfully in high stakes testing situations, such as screening applicants for military accession. An undependable applicant knows how to respond to a personality item to appear dependable and the screening organization typically has no other information on the applicant to dispute the applicant's assertion that the he or she is dependable. This is the sad state of affairs in measuring personality and temperament in high stakes testing situations.

A variety of methods have been explored to measure personality and temperament in a faking-resistant manner. Some methods work better than others. Some methods seem to work initially (e.g., forced-choice formats) but later replications of the methods fail. To date, there is

no method that is widely accepted as an effective way of measuring personality and temperament in a faking-resistant manner.

This project explores the feasibility of using situational judgment tests to assess personality and temperament such that applicants cannot improve their scores through faking. A situational judgment test item presents a scenario describing a situation and is followed by a number of possible responses to the situation. A sample item is shown in Figure 1.

> Everyone in your work group has received a new computer except you. What would you do?
>
> A. Assume it was a mistake and speak to your supervisor.
> B. Ask your supervisor why you are being treated unfairly.
> C. Take a new computer from a co-worker's desk.
> D. Complain to Human Resources.
> E. Quit.

Figure 1. Illustrative situational judgment item.

One could ask a respondent to pick the response that the applicant would most likely do in this situation. Such a response instruction is called a behavioral tendency instruction because it asks for a self-report of typical performance. Personality and temperament tests also use behavioral tendency instructions because the tests ask whether the statement (e.g., I am dependable) is a true or not a true description of themselves. A truthful respondent will give a truthful response. A lying respondent will give what they believe to be the best response. Behavioral tendency instructions are easily faked.

One could alter the response instruction for the situational judgment test to request that the respondent identify the best (most effective) response. This is a request for knowledge and not for behavioral tendency. A respondent seeking to lie to score well on the test will give what they believe to be the best response. A truthful applicant also will give what they believe to be the best response. With knowledge instructions, one turns a situational judgment test into a knowledge test where all applicants, both truthful respondents and those seeking to lie, are behaving in the same way. Both types of applicants are trying to offer the best answer. There can be no faking.

Knowledge instructions on a situational judgment test make the test a knowledge test the same way the Armed Services Vocational Aptitude Battery (ASVAB) or other aptitude tests measure mathematics knowledge. Consider this item: What is the cube root of 4,913? This is a knowledge item. One either knows the answer or one does not. Both truthful applicants and those who seek to lie to get a good score have the same goal, which is to provide the best answer to the question. With this mathematics item, as with a situational judgment item with knowledge instructions, one can report the correct answer, or one can guess at the answer, but one cannot fake the answer.

Measuring personality and temperament with a knowledge test is admittedly an odd thing to do. However, there is research support for this idea. First, there is research evidence that situational judgment tests with knowledge instructions do assess personality and temperament. Second, there is research evidence that one cannot improve one's score through faking on a situational judgment test with knowledge instructions.

Research Approach

Volunteers complete a situational judgment test, a cognitive abilities test (an ASVAB surrogate), and a forced-choice personality test (an AIM surrogate). They also complete a self-report measure of their tendencies to quit activities or to fail to complete their obligations. In addition, they indicate whether they are high school graduates, non-high school graduates, or GED holders. They also indicate whether they have served in the military, and, if so, whether they completed their first term of duty before being discharged. We seek to examine whether the situational judgment test can predict attrition-related variables over and above the cognitive ability test and the forced-choice personality test.

Accomplishments

The data collection instruments were built and processed through the human subjects review. Data collection began in the last week of October 2007.

Contributions to Basic Science

This research is expected to advance basic science in several ways. The research will evaluate the extent to which:

- Personality and temperament can be measured in a faking-resistant manner through a measurement of knowledge,
- Such measurement can result in readily interpretable personality and temperament scales, and
- Raters can predict the ability of item responses to measure personality/temperament traits as well as attrition-related variables.

The expected products of this research are:

- A situational judgment test that assesses personality and temperament in a faking-resistant manner, and
- A method to evaluate the correlates of an item prior to its administration.

Potential Army/Military Applications

The military has a great cognitive abilities test (the ASVAB). What the military needs is a better way to measure job-related and attrition-related personality and temperament. Some

instruments have shown potential in this regard. We anticipate that a situational judgment test with knowledge instructions will prove to be an effective tool in measuring job-related and attrition-related performance.

Future Plans

This three-year research project is on target to meet its goals. In the first seven months of the project, we developed the instruments and cleared them through human subjects review. Currently, we are collecting data. We have two data collection sites up and running and anticipate having three sites operational during November. The rest of the research program is devoted to data collection.

Multimedia Assessment of Emotional Abilities:
Development and Validation

Contract #: W91WAW-07-C-0025   **Contract Dates:** 05/14/2007 to 05/13/2009
Institution: Educational Testing Service   **PIs:** Richard D. Roberts and Ralf Schulze

Problem(s)/Research Question(s): Can limitations associated with existing measures of emotional abilities (EA) be overcome by developing multi-media assessments and using innovative methodologies and advanced statistical models? How are these new measures related to meaningful outcomes such as coping with stress, decision-making, teamwork, and leadership?

Technical Barrier(s): Most current measures of EA use self-report techniques, which overlap with existing personality measures and are easily faked. Performance-based tests of EA are more promising, but existing measures are limited. These measures have problematic scoring keys, lack ecological validity, and have questionable psychometric properties.

Significance/Impact for Basic Research:

- The proposed research will result in the development of a taxonomy of EA and a set of methods for measuring them. These methods could be used for assessing EA, either for personnel selection and classification or for assessing progress in EA training.

- Understanding the role of EA in relation to other psychological constructs will provide a more comprehensive theory of human capability and performance.

- Cognitive abilities have proven valuable in predicting a range of outcomes valued by society; this research will document the incremental predictive validity afforded by EA and will identify outcomes that are particularly related to emotional abilities.

Potential Transitions – Knowledge gained by this research and development effort may be useful for:

- ARI's applied selection research programs (e.g., Selection, Classification, & Performance Metrics for the Future Force Soldier; Performance Measures for 21st Century Soldier Assessment);

- ARI's applied training research programs (e.g., Future Force Warrior Training). Multimedia assessments may be retooled for this purpose and could, in principle, be a type of formative assessment;

- ARI's leader development programs (e.g., Developing Leaders in A Changing Army, Assessing and Developing Leaders for Special Operations Units).

Overview

In order to make better decisions, are some people better at using and managing their emotions, as well as interpreting and understanding the emotions of others? These are the

essential questions of research into emotional abilities (EA, also referred to as emotional intelligence). This project is concerned with how we can determine if some people are better in using their own emotions and understanding others. The project also is concerned with developing methods that reliably and validly measure differences in the ability to interpret, use, understand, and manage emotions, both in oneself and in others.

The importance of this research becomes clear in light of the purported outcomes associated with elevated EA (see Mayer, Roberts, & Barsade, 2008; Zeidner, Roberts, & Matthews, in press). Individuals with high EA are thought to make the right decisions under conditions of stress, to be better team players, to make better leaders, to refrain from damaging health behaviors, to have more fulfilling relationships; in short, to be more adaptive to their environment. The current research seeks to establish a reliable method for measuring EA and to provide validity evidence for the concept. Part of this demonstration involves relating EA to meaningful outcomes (e.g., coping with stress, leadership). Another part of this demonstration involves relating EA measures to established indices of emotion-related behaviors, intelligence, and personality, in line with theoretically predicted relationships (i.e., positive relationships with emotion and intelligence measures, weak relationships with personality, etc.).

Acknowledging limitations in extant measures of EA, a feature of this research is the development of new, multimedia (video- and text-based) test procedures to assess these emotional abilities. In addition, each measure may be classified according to a taxonomic model that we have developed in recent research (Mayer et al., 2008).

Research Approach

Tests of EA will be constructed and given to samples of community college and university students, along with measures of personality, ability, and outcomes. The methodologies used to develop the EA tests include: (1) the situational judgment test (SJT) paradigm (where participants rate a scenario for emotional relevance and/or salience); (2) an emotional principal-agent paradigm (EPAP, where event-emotion contingencies in others have to be perceived and memorized and emotion-behavior contingencies inferred from observed behavior to predict future behavior); (3) a cloze technique (where an emotional term completes, for example, a quote made by a famous philosopher); (4) various information processing paradigms (with emotions as stimuli and speed of response the variable of interest); and (5) an implicit association technique (where an individual's implicit association of emotions with words and situations are assessed).

All of the tests will be computerized, where (1) and (2) comprise video sequences as constituent components, (3) is a text-based multiple choice test akin to classical reasoning tests, and (4) and (5) include both figural and auditory (reaction time) tests. The use of multiple media is essential for several reasons. First, it enables the assessment of certain emotional abilities that cannot easily be assessed with other methodologies. Second, it broadens the conceptualization of EA beyond tasks that draw excessively on verbal ability, a problem that has plagued the field to date. Lastly, it serves to enhance the ecological validity of the test paradigms.

The proposed efforts will evaluate the psychometric properties of these new measures, including tests for measurement invariance (i.e., is the test measuring the same things in different

sub-samples) and examination of sub-group differences (e.g., ethnic groups) at the latent variable level. Estimation of structural relations between latent variables underlying both the new and established (EA) measures will serve to test the proposed taxonomic model and provide evidence for the new measures' convergent and discriminant validity. Further validity evidence will be collected by estimating the new tests' relationships with outcome measures (e.g., GPA, coping with stress, social support) where intelligence and personality variables are controlled for.

If data are supportive of the new measures' quality, in the final year of funding we will conduct an experiment on the relations of EA to stress. Also in the final year of funding, an effort examining a broader criterion space, including leadership and psychological well-being (especially relations to happiness, as assessed by the Day Reconstruction Method) will be conducted. One or two of these efforts also would examine additional psychometric properties of a selection of these EA instruments (e.g., test-retest reliability).

Accomplishments

A review of the literature has indicated that the proposed methodologies can be adapted for use in measuring EA (Orchard et al., in press). In addition to this activity, we have obtained critical incidents from a number of sources that we have furnished to assessment development specialists, who have begun working on script writing. A brief example of one such item, measuring facets of EA, is shown in Figure 1.

Establishing Scenario (video: 60-90 sec)

Professor (to student, with moderate concern): "You've been doing well on your assignments, but you really need to work on your exam performance. You scored only a 72 on the last test."
Student (to professor, sarcastically): "Well, maybe that's because the topics I think up for my papers are more interesting than your exam questions."

Problem-Solving Item (video: 60-90 sec)

How should the professor respond? [format could be multiple-choice, rank-order, Likert-scale]
    (A) *(snidely)* "Not if you're the one reading the papers."
    (B) *(off-handedly)* "More interesting, perhaps—more illuminating, no."
    (C) *(assertively)* "I know how to make a good exam: I'm an Outside Item Writer for ETS."
    (D) *(helpfully)* "I understand that you're a bit frustrated, but if you want we can discuss ways
        to prepare for the tests."

Rationale Item (text)

What are your reasons for choosing ([examinee's answer])? Select all that apply.
    [ ] The professor uses humor to diffuse the situation.
    [ ] The professor makes an important distinction between assignments and tests.
    [ ] The professor establishes her qualifications and credibility.
    [ ] The professor shows an interest in helping the student improve.

Figure 1. Sample script related to the multimedia SJT.

We have developed or collected the vast majority of the additional measures (e.g., personality, cognitive ability, and outcome measures) required as part of the multivariate design. Further, we have met regularly with the film production company (Creative Media for Learning) and are in the process of finalizing logistical details surrounding the film shoot (to be held in early 2008). We also have built prototypes that will inform programming of the multimedia SJT and EPAP. One of these is shown in Figure 1.

In addition, we have continued longstanding active collaborations with leading researchers, and have collected data that will allow us to refine the proposed multivariate design. In one such activity, we have conducted meta-analysis of self-report and performance-based measures, highlighting major limitations with the former technique in particular. Based on these data, we feel fairly confident that existing self-report measures are merely proxies for personality and should not be considered measures of EA (Roberts et al., in press). In another activity, conducted with collaborators at the University of Berkeley, we have used item response theory to resolve some identified psychometric problems associated with consensus-based measurement of EA. This later effort is currently being prepared for publication in a peer-review journal.
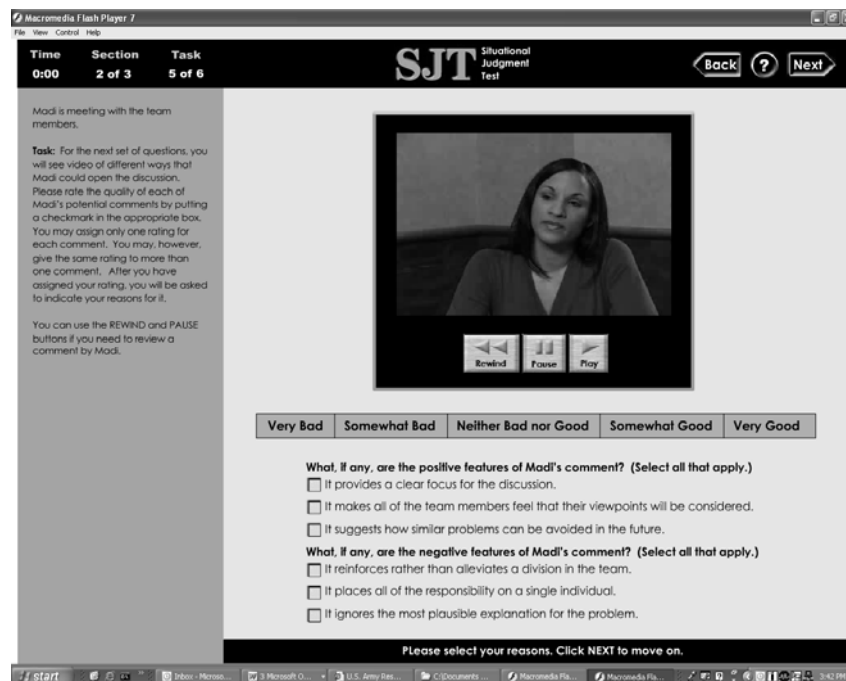

Figure 2. Prototype of interface for the multimedia SJT.

We also have completed a draft of a book on EA (Zeidner et al., in press) and a journal article on EA that looks set to be published in the American Psychological Association (APA) journal *Emotion*. These artifacts bear mentioning here because they emphasize problems, pitfalls, and promises associated with the measurement of EA. They also give a detailed description of potential methods of resolving measurement issues and lay a theoretical platform for relations between EA and outcomes of practical significance, which this large-scale multivariate effort can then resolve.

Contributions to Basic Science

The EA measures under development use a variety of different methodologies: situational judgment, implicit association, information processing, reasoning, and principal agent paradigm approaches. Current research indicates that each of these methodologies offers a new method of assessing EA, in a more objective and valid way.

The research we are conducting will provide insight into the domain of EA, potentially providing an empirical taxonomy for placing tasks within, allowing more research in the area, and informing our understanding of how humans use emotions in making decisions. Linking these EA assessments to other domains, such as teamwork and leadership also holds promise of more compelling models of human performance. Additionally, being able to differentiate between those who are able to appropriately recognize, manage, and use their emotions (especially in highly charged emotional situations) will enable organizations to recognize which individuals are likely to perform well in emotion-laden situations.

Potential Army/Military Applications

The research we are conducting can be used in the ARI's applied selection research programs, with an eventual segue into use for classification, selection, placement, and promotion of personnel. Additionally, information regarding EA may be useful in providing training to existing personnel. In particular, certain branches (and the MOSs within those branches) may find that using EA as an assessment tool for the purposes mentioned above can increase Soldier success. As an example, ARI may find that Psychological Operations, Adjutant General's Corps, and Military Intelligence personnel with higher EA perform more efficiently and to a higher standard than those with lower EA. The relationship between higher EA and performance of duties required by these branches is not known at this time, but for personnel needing to make evaluations of intent and gauging emotional or social situations, such as might be encountered by these personnel, it is likely that higher EA will provide substantial benefits. Additionally, personnel in command positions may benefit from higher EA. Our research enables us to first develop a measure of EA and then validate that measure, such that its relationship with valuable outcomes is known.

As a training tool, a measure of EA could be used to identify personnel that need increased understanding of how to work with emotional situations, and then provide training to those individuals so they are able to better work with, understand, and manage their own and others' emotions. The developed SJTs also might serve as training tools in their own right, with the potential for these to serve a formative assessment function.

A presumed advantage of developing these EA assessments relies on their purported freedom from biases seen in traditional cognitive measures (i.e., they hold promise of correcting adverse impact). In addition, these measures might provide potential predictive information regarding a range of behaviors, which cognitive measures do not cover. These features may make these attractive noncognitive supplements to future versions of the ASVAB as recommended by Drasgow et al. (2006).

Future Plans

During 2008, tests will be developed, filmed in a multimedia format, and programmed into a single delivery system. Additionally, a usability analysis will support any necessary modifications to this platform and identify human factors issues. A pilot test of the measures will be conducted to provide an initial evaluation of the measures. Then, a multivariate designed effort (N=1000) will be conducted, with data analysis completed by 1/13/09.

References

Drasgow, F., Embretson, S. E., Kyllonen, P. C., & Schmitt, N. (2006). *Technical Review of the Armed Services Vocational Aptitude Battery (ASVAB).* Technical Report FR 06-25. Alexandria, VA: HumRRO.

Mayer, J., Roberts, R. D., & Barsade, S. G. (2008). Human abilities: Emotional intelligence. *Annual Review of Psychology, 59*, 507-536.

Orchard, B., MacCann, C., Schulze, R., Matthews, G., Zeidner, M., & Roberts, R. D. (in press). New directions and alternative approaches to the measurement of emotional intelligence. In D. Saklofske, J. Parker, & C. Stough (Eds.), *Measuring emotional intelligence.* New York: Springer.

Roberts, R. D., Schulze, R., & MacCann, C. (in press). The measurement of emotional intelligence: A decade of progress? In G. Boyle, G. Matthews, & D. Saklofske (Eds), *Sage personality handbook series*. Thousand Oaks, CA: Sage Publications, Inc.

Zeidner, M., Roberts, R. D., & Matthews, G. (in press). *Emotional intelligence primer.* Cambridge, MA: MIT Press.

Hispanics and African Americans in the U.S. Military: Trends in Representation

Contract #: W74V8H-05-K-0007
Institution: University of Maryland

Contract Dates: 11/01/2004 to 10/31/2008
PI: David R. Segal
Co-PIs: Mady Wechsler Segal and
Meyer Kestnbaum

Problem(s)/Research Question(s) – What are the trends in ethnic, racial, and gender representation in the military, and how do these vary among services? How do economic market factors versus cultural factors affect these trends?

Technical Barrier(s) – Most data do not distinguish among different Hispanic nationalities, but standard measures of race and ethnicity have changed. There are variations in estimates of Hispanic representation in the American youth population and in estimates of the proportion of that population that is eligible to serve in the military.
.
Significance/Impact for Basic Research – These analyses address the relationship between attitudes (propensity to serve) and behavior. This analysis in representation trends suggests that labor market dynamics have more of an impact on enlistment patterns than culture.

Potential Transitions – Knowledge gained in this basic research may be employed in:

- Development of recruitment strategies and campaigns,
- Informing personnel management policies (i.e., reenlistment incentives), and
- Providing a general understanding of the military's demographic composition.

Overview

The transition from a conscription-based force to an all-volunteer force (AVF) brought issues of social class, race/ethnicity, and gender to the forefront of military recruitment and retention considerations. During the transition planning, social analysts anticipated an overrepresentation of those populations disadvantaged in the labor force, and in particular of racial minorities from the lower socioeconomic statuses. Issues of gender were not yet raised because planners did not anticipate increases in women's military participation, although the increased presence of women contributed to the success of the AVF.

The purpose of this research is to document the current representation of Hispanic and African American men and women in the armed forces in the context of trends in propensity to serve. Our research suggests that labor market dynamics have more of an impact on enlistment patterns than culture. This research also discusses the explanations and implications for these trends and anticipates future military demographics.

Research Approach

The data for this research consists of survey data on high school seniors as well as information collected by the Department of Defense. We compared trends in propensity to serve and enlistment. We further analyzed this data by focusing on the intersections of different social diversity characteristics, with an emphasis on minority characteristics (i.e., Hispanic female or African American male).

Accomplishments

African Americans have been overrepresented in the U.S. military, especially the Army, since the start of the AVF in 1973, and they have been dramatically overrepresented among military women. As Figures 1-4 show, propensity to serve has been declining among African American and Hispanic men and women, particularly since 1990.
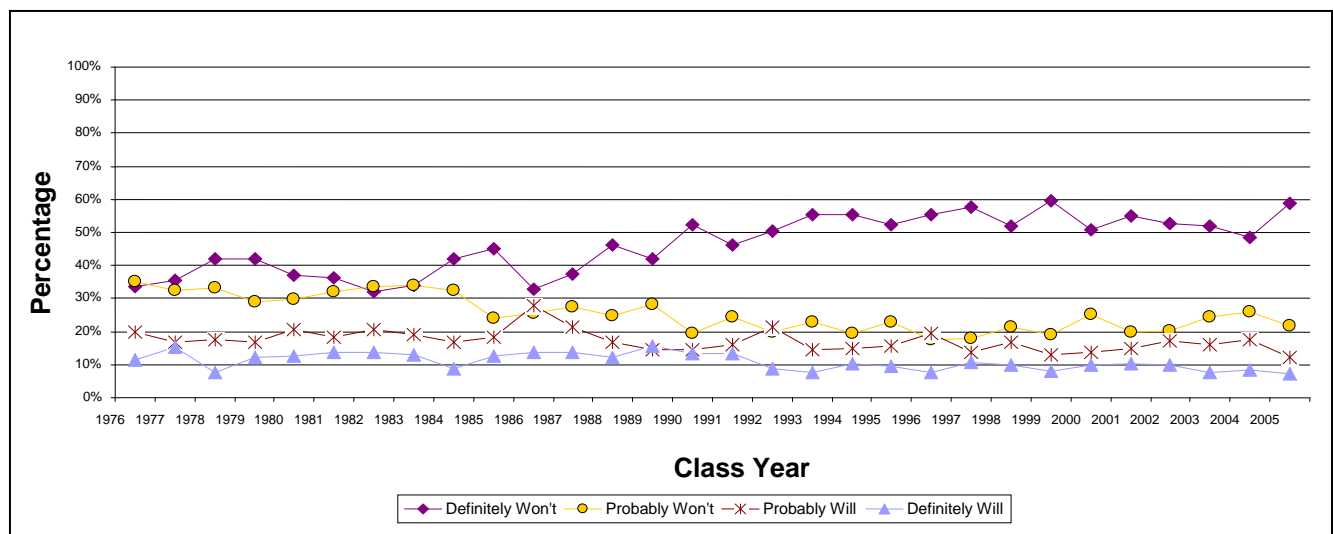
Figure 1. Trends in propensity to enter the armed forces: high school seniors, 1976-2005, Hispanic males.

Note. Sources for Figures 1-4: Bachman, Jerald G., Johnston, Lloyd D., & O'Malley, Patrick M.. *Monitoring The Future: A Continuing Study Of The Lifestyles And Values Of Youth*, 1976-2005 [Computer file]. Conducted by University of Michigan, Survey Research Center. ICPSR07927-v4. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor].
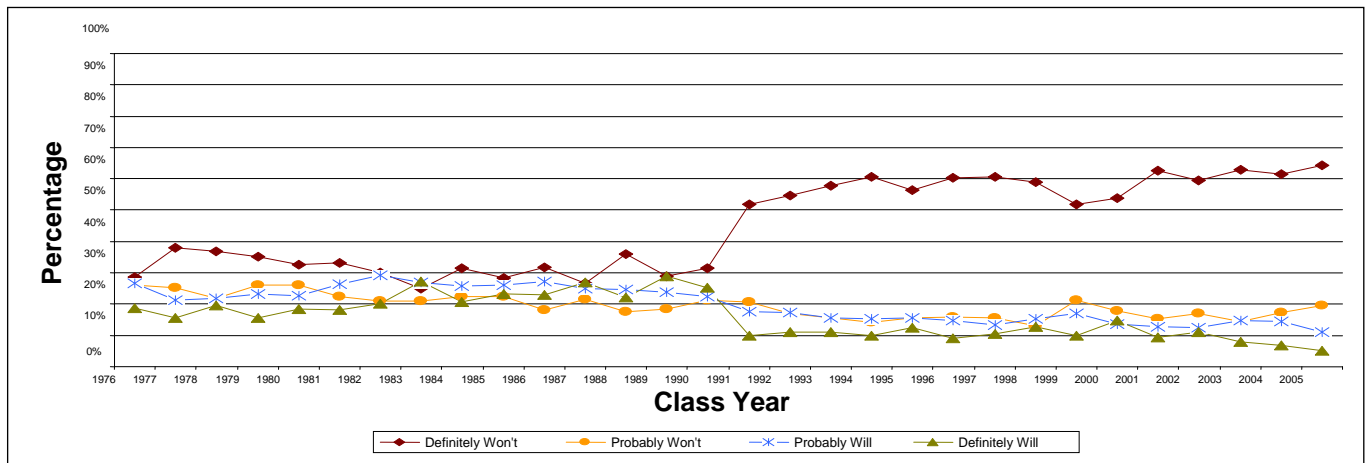
Figure 2. Trends in propensity to enter the armed forces: high school seniors, 1976-2005, Black males.
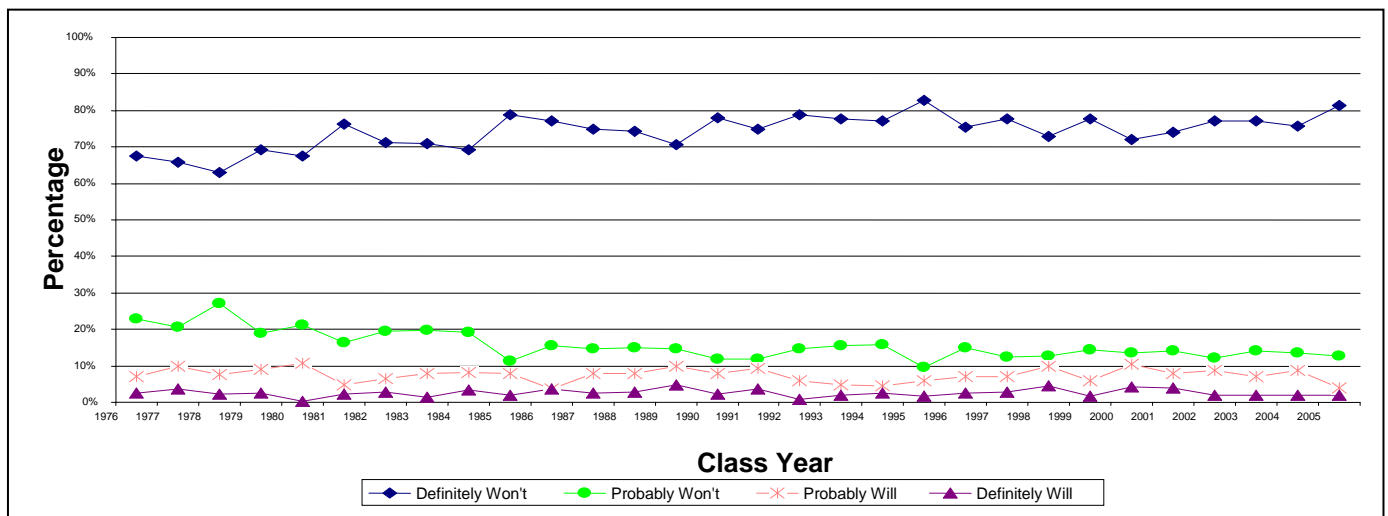

Figure 3. Trends in propensity to enter the armed forces: high school seniors, 1976-2005, Hispanic females.
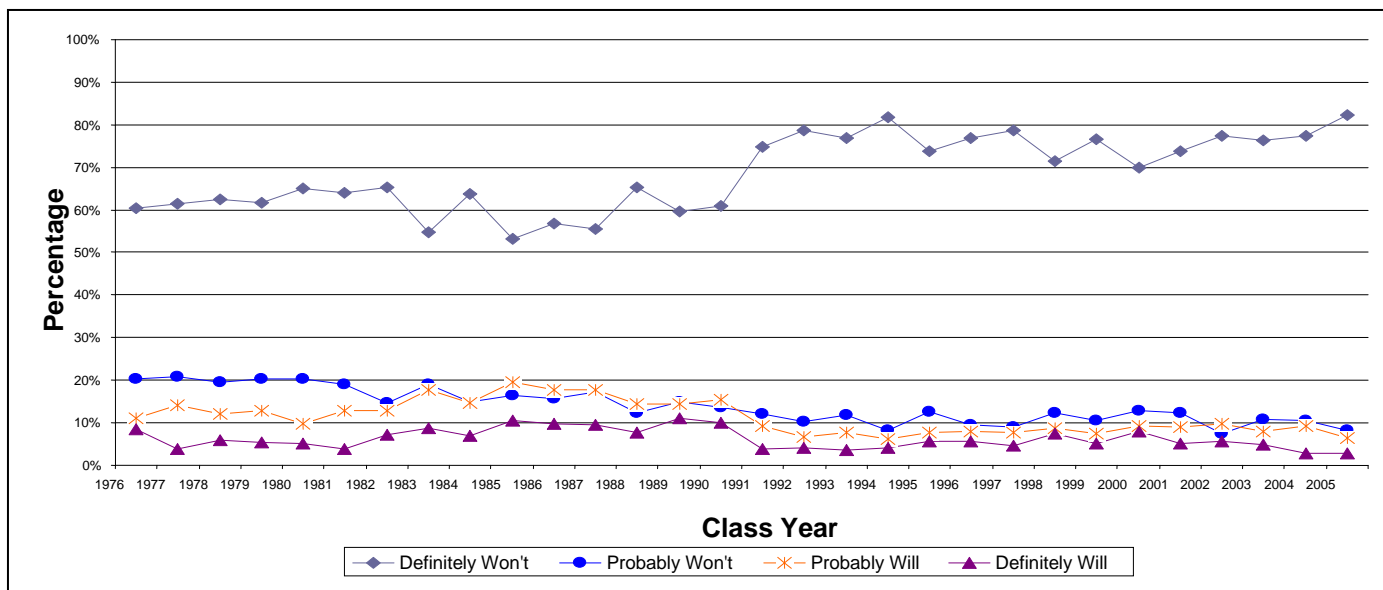
Figure 4. Trends in propensity to enter the armed forces: high school seniors, 1976-2005, Black females.

Recently, as Figure 5 shows, accessions of African Americans have likewise declined.

Figure 5. Blacks as a percentage of active duty enlisted personnel, by service, with civilian labor force (CLF) aged 18-44, FYs 1973-September 2006.

Note. Source for Figures 5 and 6: Department of Defense, *Population Representation in the Military Services, Fiscal Year 2002*, March 2004, pp. 3-5; Department of Defense, *Population Representation in the Military Services, Fiscal Year 2004*, May 2006, pp. 3-5; September 2006 data provided by the Office of the Deputy Under Secretary of Defense (Equal Opportunity).

At the same time, Hispanics, who constitute a growing segment of the U.S. population and have been underrepresented in the military, especially among the officer corps, have increased their enlistment, despite their declining propensity to serve, as shown in Figure 6.

Moreover, despite the traditionally masculine culture of the military institution, and the gendered nature of Hispanic culture, Hispanic women now comprise a larger percentage of female personnel than Hispanic men of the male military population.
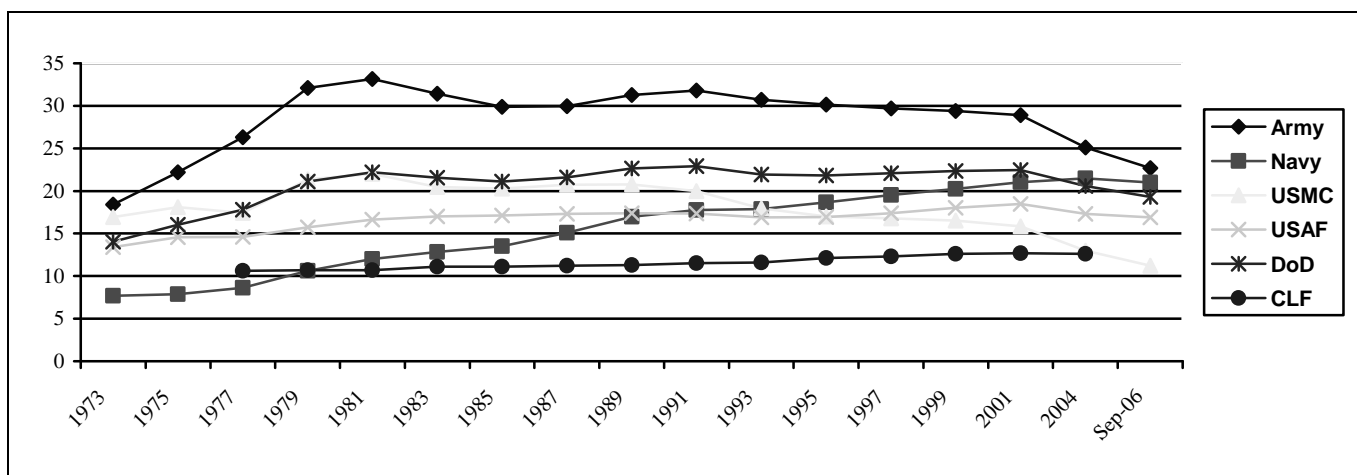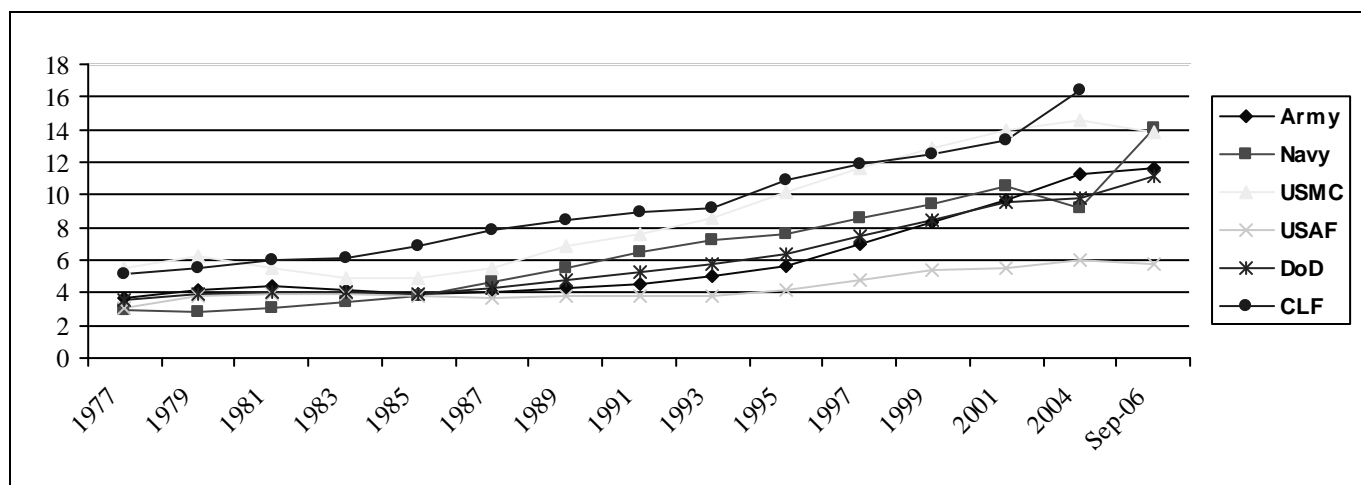

Figure 6. Hispanics as a percentage of active duty enlisted personnel, by service, with civilian labor force (CLF) aged 18-44, FYs 1977-September 2006.

Contributions to Basic Research

Early social scientific projections of African Americans being overrepresented in the volunteer military were correct, and limitations on the service of women had to be lifted to make the volunteer force succeed. Had it not been for an overrepresentation of African Americans, and an increasing proportion of women in uniform, the recruiting goals of the volunteer force would not have been met. The intersection of race, class, and gender became crucial for recruiting the volunteer force and remain important for recruiting and retention initiatives today. More recently, the growth of the Hispanic population, particularly in the lower socio-economic strata, has required the inclusion of ethnicity at this intersection. This research moves the analysis of military recruiting in this direction. We conclude that the overrepresentation in the armed forces of high school educated African Americans (particularly women), and the increase in the enlistment of Hispanics, including women, speaks to the impact of class, gender, and race/ethnicity on the labor market dynamics that produce our volunteer military. It appears that the military services are making up for declines in African Americans joining the military in part by recruiting more Hispanics. If the current recruiting and reenlistment trends continue, we expect that representation of Hispanic men and women in the U.S. military services will continue to increase, while African American representation will decline even further.

Potential Army/Military Applications

This research demonstrates how labor market dynamics continue to exert more influence than cultural factors in explaining patterns of military service. Specifically, the reasons cited by racial/ethnic minorities for joining the military were more economic than those cited by White personnel. However, this distinction does not account for the current rates that demonstrate that

116

Hispanic men and women are more highly represented among enlisted personnel in the Marine Corps than the other services. Thus, this research demonstrates the overall importance of economic reasons regarding enlistment decisions, but also demonstrates the relevance of cultural factors specific to each service.

This research may further inform both overall recruitment strategy within the Department of Defense and the military services as well as shape service-specific tactics for recruitment and retention. The challenges of manning an AVF require leaders and policymakers to continuously reevaluate labor market demographics, particularly in regard to how the market interacts with social diversity characteristics.

Future Plans

Areas of future research include exploring the role of culture within enlistment decisions as well as examining why the correlation between propensity to serve and enlistment does not hold in the Hispanic context. The research will be expanded to other ethnic groups such as Native Americans, and the linkage between attitudes (propensity) and behavior (enlistment) will be explored further.

Temporal Investigations Into the Relationship Between Affect and Discretionary Work Behavior

Problem(s)/Research Question(s) – What is the nature of the structure of citizenship and counterproductive work behavior (collectively known as discretionary behavior)? How does discretionary behavior vary over time within individuals, and does its expression vary as a function of changes in persons' affective (emotional) states? Also, what are some of the variables that would alter the relationship between emotions and discretionary behavior at work (e.g., leadership, climates)?

Technical Barrier(s) – Models of how discretionary behavior changes within individuals and the effects of emotions on discretionary behavior had not previously been well developed. This lacuna existed in large part because the methodological and statistical tools to study such phenomena have not existed until recently.

Significance/Impact for Basic Research – Understanding the structure of discretionary behavior and emotion-behavior linkages leads to more comprehensive models of human performance and a focus on *within*-person changes in performance over time.

Potential Transitions – Knowledge gained in this basic research may be usefully employed in:

- WP 215 — Objective Force Warrior Training,
- WP 283 — Validating Future Force Performance Measures,
- WP 302 — Ground Systems Team Training, and
- WP 333 — Influence Strategies for Leaders.

Overview

In spite of the recent interest in citizenship and counterproductive work behavior, and their documented importance to organizational functioning, much remains to be discovered about these criteria. Our program of research integrates the burgeoning literature on the immediate consequences of momentary affective states (see Brief & Weiss, 2002, for a review of this literature) with the literature on discretionary work behavior. We conducted several research efforts that explore and map individuals' naturally-occurring on-the-job citizenship and counterproductive behavior, the emotional antecedents of such behavior, and the variables that potentially influence emotion-behavior relationships (e.g., leadership, climates).

Our project had four main goals. First, we examined the structure of work behavior by determining behavioral co-occurrence and switching within persons over time. The main aims were to determine which behaviors occur together on a given time occasion (i.e., behavioral co-

occurrence) and to predict changes from one behavioral state on one occasion to the next behavioral state on a subsequent occasion (i.e., behavioral switching). The extant research literature had not examined the possibility that the dynamic relationship (i.e., using "time" as the level of analysis) between citizenship and counterproductive behavior may not mirror the static relationship (i.e., using "person" as the level of analysis)—an oversight that needed to be rectified. The structure of discretionary behavior also needed to be established before we explore its emotional antecedents.

Another related focus involved establishing employees' immediate emotional states as the causes of citizenship and counterproductive work behavior. Research examining the impact of emotional states on citizenship behavior and, particularly, counterproductive behavior had been very meager. Although basic research on emotions suggests that momentary affective states are an important influence on both citizenship behavior and counterproductive behavior, a thorough and overarching framework for studying these relationships in detail was needed. We used Affective Events Theory (Weiss & Cropanzano, 1996) as a unifying framework for understanding these relationships in detail. A fundamental proposition of our research was that changes in the likelihood of engaging in citizenship and counterproductive behavior over time are related to changes in a person's affective states.

The experience of negative affective states does not always lead to counterproductive work behavior; similarly, citizenship behavior does not always follow positive affective states. People's ability to "influence which emotions they have, when they have them, and how they experience and express those emotions" should influence whether or not a given behavior is exhibited (Gross, 1999, p. 557). This process, known as emotional regulation, needed to be integrated into the study of the way emotional states instigate counterproductive behavior. Our research assessed whether the ability to regulate one's emotions at any given time influenced the expression of counterproductive and citizenship behavior. By investigating the role of emotional regulation, researchers will ultimately be able to develop taxonomies of regulatory strategies that can be used to control the dysfunctional behavioral consequences of emotional states.

Employees' emotions and behavior occur in a work context that can play an important role in facilitating or inhibiting certain behaviors following an emotion-instigating event. Therefore, we examine leaders' ability to regulate their subordinates' behavior via the creation of "climates," leaders' adeptness at identifying others' emotions, and their skill at managing and altering these emotions in an effort to avert counterproductive behavior on the part of the employee.

Research Approach

Contrary to traditional cross-sectional studies that look at differences in key variables between individuals, we adopted a within-person approach that examined how variables change within an individual over time. Given that the structure of discretionary behavior and the associations between affect and such behavior are temporally sensitive, we used a dynamic data collection approach known as Ecological Momentary Assessment (EMA). Ecological momentary assessment has, of late, begun to gain popularity in the organizational psychology literature (Weiss, Nicholas, & Daus, 1999; Alliger & Williams, 1993) because it allows for the

quantitative analysis of work behavior in its natural context (Hormuth, 1986)—not only its natural physical context but also its natural temporal context. EMA can therefore be used to examine issues concerning the relatedness of variables within persons over time.

In order to capture phenomena as they occurred, we asked participants to complete multiple daily questionnaires on Personal Digital Assistants (PDAs), such as Palm Pilots™, for data collection. The durations of the experiments varied as a function of the particular samples used and the research questions of interest; however, experiments typically lasted 10-15 working days. Individuals were asked to provide reports of their discretionary behavior and affective states several times per workday.

Accomplishments

We have completed data collection on all four experiments. Experiment 1 was presented at the 2006 annual meeting of the Society for Industrial and Organizational Psychology, as part of a symposium on innovative methodological research on citizenship and counterproductive work behavior. Subsequently, Experiment 1 was submitted to the *Academy of Management Journal* (one of the premier outlets for Industrial and Organizational (IO) psychology research). Due to the novelty of our research approach and the need to familiarize the editor and reviewers with this approach, the review process at this journal has taken a longer time than expected; however, we are hopeful that the paper will be accepted soon. Experiment 2 and the pilot data from Experiment 4 were presented at the 2007 annual meeting of the Society for Industrial and Organizational Psychology. We have submitted Experiment 3 to the 2008 annual meeting of the same conference. We also have submitted our final technical report to ARI.

Contributions to Basic Science

Contrary to most traditional models of work performance, our focus on employees' affective states as predictors of performance extends the existing literature by providing a framework for understanding how emotions influence discretionary work behavior over time. Moreover, our projects integrate theories of affective experiences and discretionary work behavior. Thus, our research helps to re-focus the organizational sciences on the study of affect. In addition, our research examines how people experiencing negative affective states attempt to minimize the occurrence of dysfunctional work behavior through the use of emotional regulation strategies.

Potential Army/Military Applications

Negligent and willful counterproductive/deviant behaviors are especially of great importance in the military, where the cost of such behavior during combat operations can often be measured in terms of lost lives: those of civilians, of one's fellow Soldiers, and—in cases of excessive brutality and wanton disregard for norms of civilized combat—even of enemy combatants and detainees. Consequently, it is no great surprise that the United States military has long recognized the importance of studying counterproductive (also called unethical or deviant or delinquent) behavior among its own Soldiers (e.g., Bell & Holz, 1975; Lennon, 1994). Our research is unmistakably relevant to the U.S. Army in its quest for the "good Soldier."

The major implication of our research is that behavior and affect are transient in nature, and that citizenship and counterproductive behavior are largely driven by affect. In contrast to most research and theorizing on unethical behavior, which focus on cognitive antecedents, the present research strongly suggests that the antecedents to at least some forms of unethical behavior are in fact affective—and hence dynamic or temporally-based. These behaviors may therefore have a dynamic, temporal structure that is not identical to their static structure. We suggest first determining the temporal structure of discretionary behavior in civilian settings with their inherent lower-intensity stressors, and then conducting more applied research with the military to examine the extent to which previous findings hold up in more extreme situations— for example, patterns of Soldier behavior under (simulated) enemy fire on the battlefield. This subsequent applied military research can take the form not only of analysis of emotion-behavior and behavior-behavior links in simulated real-time battlefield environments, but also of retrospective investigations into egregious incidents committed by Soldiers (e.g., torturing prisoners, executing civilians, shooting their own colleagues, committing mutiny against superior officers, etc.), and their affective antecedents. Such research would provide valuable "red flags" that, if heeded, could head off extremely serious incidents before they actually occur.

Another implication of our proposed research is that affective responses to stressful and/or aversive events can be regulated, preventing expression of counterproductive behavior. Emotional regulation strategies can be executed either by the Soldier himself or herself, or by the commanding officer. Research on emotion-behavior links and emotional regulation are perhaps even more important in the military than in other settings, given research suggesting that verbal behavior in the military has ambiguous emotional nuances, making correct interpretations of the speaker's intent especially difficult (Wirshbo, 1990). Officers unable to correctly discern their subordinates' emotional states are unlikely to be able to respond with appropriate emotional regulation strategies. Future research in military settings should therefore examine situational and individual differences in emotional judgment related to oneself or others.

Additionally, our research can lay the groundwork for the study of officer-sponsored "climates" that create contingencies between certain behaviors and certain outcomes, thereby attenuating emotion-behavior linkages. For instance, an officer who clearly articulates and consistently enforces a policy of sanctions for unethical behavior or poor disciplinary practices may be able to reduce the incidence of such practices regardless of the frequency and magnitude of stressors and consequent aversive reactions experienced by his or her subordinates. The findings of our research efforts are likely to provide material relevant to courses in leader development and training. In particular, a within-subjects approach may be more useful than a between-subjects approach in terms of providing feedback to leaders.

Final Summary

The role of affect (mood and emotions) in determining organizational behavior has only recently begun to be appreciated. In part, this is due to the fact that affect is highly volatile within persons—for example, mood may change from hour to hour—and that novel conceptual, methodological, and analytical approaches are needed to study affect-behavior linkages. Unlike most studies, where participants are surveyed once or at most twice, the Ecological Momentary

Assessment approach used here required each participant to be surveyed multiple times per workday for 2-3 workweeks. This approach allowed us to study how affect and behavior varied within persons over time.

One of the major findings was that workplace behaviors like organizational citizenship behavior and counterproductive work behavior, which were previously thought to be stable over time within persons, are in fact highly volatile. Demonstrating within-person volatility is a prerequisite to demonstrating that these behaviors are in fact consequences of mood and emotions at work. A second finding was that, within persons over time, organizational citizenship behavior and counterproductive work behavior were virtually unrelated to each other—contrary to conventional wisdom that they should be strongly negatively related. This finding suggests that it is an oversimplification to believe that a person will refrain from counterproductive behavior during time periods when he or she is engaging in citizenship behavior, or vice versa. In other words, employee performance is a lot more complex than previously believed.

A third major finding was that of consistently strong linkages between affect and work behavior within persons over time. To us, this is unsurprising. However, to a field that, despite pioneering work in the area (e.g., Hersey, 1932), turned its back on mood and emotions—and, more generally, turned its back on phenomena occurring within, rather than across, persons—this message is unanticipated and therefore bears repeating. We hope that workplace affect soon attains (or rather regains) the status of a core topic in organizational research.

A final finding pertains to the impact of experiencing counterproductive work behavior. In one experiment, we found that people who experience uncivil behavior (a form of counterproductive behavior) are more likely to experience negative emotions, and to respond with incivility and other forms of counterproductive behavior such as interpersonal withdrawal. Future research should investigate the specific conditions under which experiencing counterproductive work behavior leads to retaliation against the perpetrator versus other responses such as retaliation against an innocent third person ("kicking the dog") or the enactment of a different form of counterproductive work behavior (such as interpersonal withdrawal).

The research program discussed above has numerous practical implications for the U.S. military. One important implication is that, regardless of whether we look at differences in performance across people (Project A) or differences in performance across time within the same person (our research), overall Soldier performance is multidimensional. The importance of studying the dimension of counterproductive or deviant work behavior, in particular, has been recognized by the U.S. military. Counterproductive work behavior is likely to be of great interest to military policymakers in the present environment, given the widespread publicity attached to the incidents of detainee and civilian mistreatment during the wars in Iraq and Afghanistan, as well as the current focus on the "strategic sergeant." The present research is moreover relevant to military work on ethical behavior, because in most cases counterproductive behavior would be considered unethical, whereas citizenship behavior would be considered highly ethical.

A second implication is that job performance (or each facet thereof) is highly volatile within a given person over time. In other words, over time, there are likely to be numerous deviations—and occasionally very large deviations—from a Soldier's average level of performance. Yet, most research to date has assumed within-person stability and focused on between-person differences (i.e., individual differences), or at best has focused on relatively long-term (e.g., months or years) within-person change. However, our research demonstrates that short-term (e.g., hours or days) within-person change also exists, and is also very important. Because of the long-term and, especially, short-term changes in the level of a Soldier's job performance, it is unreasonable to expect selection tests to be able to yield Soldiers who consistently perform at very high levels. Situational factors, such as events and interventions, therefore assume greater importance. Leaders are especially important here, because they can establish climates prescribing ethical (citizenship) behavior and proscribing unethical (counterproductive) behavior.

A final implication is that, simply stated, emotions matter. In fact, they matter a great deal, because they influence work behavior and performance. It is difficult to acknowledge the importance of emotions in a traditionally masculine culture like the military. However, military policy-makers interested in "evidence-based practice" are likely to see the value in attempting to decrease both the occurrence and the impact of negative emotions among Soldiers. Indeed, there has been a history of studying emotions in the military, at least insofar as they influence behavior (e.g., Longley, 1947). In particular, the present research suggests that the precursors to (un)ethical behavior can be affective, rather than purely "rational" or cognitive. Overall, the study of emotions need not be incompatible with the Stoic philosophy that continues to influence the U.S. military (Sherman, 2005).

References

Alliger, G. M., & Williams, K. J. (1993). Using signal-contingent experience sampling methodology to study work in the field: A discussion and illustration examining task perceptions and mood. *Personnel Psychology, 46,* 525-549.

Brief, A. P., & Weiss, H. M. (2002). Organizational behavior: Affect in the workplace. *Annual Review of Psychology, 53,* 279–307.

Bell, D. B., & Holz, R. F. (1975). *Summary of ARI research on military delinquency*. (Research Report, No. 1185). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Gross, J. J. (1999). Emotion regulation: Past, present and future. *Cognition and Emotion, 13*, 551-573.

Hersey, R. B. (1932). *Workers' emotions in shop and home: A study of individual workers from the psychological and physiological standpoint*. Philadelphia: University of Pennsylvania Press.

Hormuth, S. (1986). The random sampling of experiences *in situ*. *Journal of Personality, 54,* 262-293.

Lennon, D. A. (1994). Crime in the U.S. Coast Guard, 1964-1990. *Deviant Behavior, 15*, 397-416.

Longley, A. C. (1947). Military psychology—emotions. *Military Review*, *28*, 54-59.

Sherman, N. (2005). *Stoic warriors: The ancient philosophy behind the military mind.* New York: Oxford University Press.

Weiss, H. M., & Cropanzano, R. (1996). Affective events theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work. *Research in Organizational Behavior, 19,* 1-74.

Weiss, H. M., Nicholas, J. P., & Daus, C. S. (1999). An examination of the joint effects of affective experiences and job beliefs on job satisfaction and variations in affective experiences over time. *Organizational Behavior and Human Decision Processes, 78,* 1-24.

Wirshbo, E. (1990). Can emotions be determined from words? A consideration of recent military usage. *American Behavioral Scientist, 33(3)*, 287-295.

BRU RESEARCH OBJECTIVE #4: PROVIDE FUNDAMENTAL KNOWLEDGE FOR ORGANIZATIONAL BEHAVIOR AND NETWORK SCIENCE RESEARCH

Research in this area focuses on understanding and predicting large and small group behavioral processes in dynamic social networks, whether in simulations, games, or Army organizations. This area provides support for network science collaboratory research at the Network Science Technology and Experimentation Center (NSTEC). Modeling and understanding networks underlying human knowledge systems is part of this objective.

Improved Data Extraction and Assessment for
Dynamic Network Analysis

Contract #: W91WAW-07-C-0063   **Contract Dates:** 08/07/2007 to 08/31/2012
Institution: Carnegie Mellon University   **PI:** Kathleen M. Carley

Problem(s)/Research Question(s) – The Army has a need for rapid and accurate analysis of unformatted data in order to support various operations. Examples of raw texts include open source, human intelligence (humint), and situation reports such as those entered in Command Post of the Future (CPOF). For example, there is a need to identify texts and extract from them information about various actors and associated political, economic, military, social, and infrastructure related activities. Key questions include: How can text data be rapidly processed to extract meaning and not just key entities and content analysis?; How can links be extracted?; and how can the attributes of extracted entities be identified?

Technical Barrier(s) – Dynamic network analysis tools can be used to process unformatted data. However, these tools need to be improved so that the time spent by the user in locating, ingesting, and coding data is minimized and so that analyses can be quickly upgraded as new information is collected. Entity extraction, particularly as it applies to names and locations, can be done quite well. Link extraction and inference is not currently possible. Barriers to rapid development here include the lack of adequate tagged data for system testing, and lack of appropriate unsupervised learning algorithms.

Significance/Impact for Basic Research – We propose to augment the DyNet tool suite so that a more flexible, automated, user-friendly, rapid, and accurate set of techniques are available for data extraction, coding, and assessment with particular focus on link extraction, attribute extraction, and semi-automated ontology creation. Of particular import are the tools for automated thesauri construction, anaphora resolution, list resolution, and hierarchical ontologies for second order thesauri classification. We will focus particularly on raw texts, combining text based and alternative database and Excel file information, and data that is dynamic, continually being augmented, or is time stamped. The proposed tools will support the DOD need to rapidly assess socio-cultural and political information, rapidly identify key influences and opinion leaders and changes in their opinions, and enable rapid utilization of data collected by diverse subject matter experts for modeling and forecasting the impact of alternative courses of action.

Potential Transitions – AutoMap, which we are extending under this project, is currently an underlying system in VIBES, which is being developed for the G2. This will be reviewed for incorporation in MAP HT. This would serve three key purposes: identification of hot topics, assessment of trends in attitudes, and extraction of meta-networks to support DIME/PMESII modeling and social network analysis. I was just approached by DeepMile, an intel firm, with the possibility of having their tool called AutoMap. AutoMap also is being explored as a tool to support counter-narcotics work for the National Guard.

New extensions to AutoMap developed for this project will be taught in courses at Carnegie Mellon University (CMU) and also at the Center for Computational Analysis of Social and Organizational Systems(CASOS) summer institute. We have been asked to teach members

of the National Guard how to use AutoMap in the counter-narcotics area, and will also be conducting a training module at the United States Military Academy (USMA).

Overview

Dynamic network analysis (DNA) is an emergent field centered on the collection, analysis, understanding, and prediction of dynamic relations (such as who talks to whom and who knows what) and the impact of such dynamics on individual and group behavior. DNA facilitates reasoning about real groups as complex dynamic systems that evolve over time. Within this field, computational techniques, such as machine learning and artificial intelligence, are combined with traditional graph and social network theory, and empirical research on human behavior, groups, organizations, and societies to develop and test tools and theories of relational enabled and constrained action.

A key aspect of DNA is the extraction of network data from texts. This involves the extraction of who, what, where, how, why, and when and the relations among given textual information that may contain a variety of typographical and content errors, translation mistakes, alternative wordings for the same event, aliases, etc. We propose to advance the state of text analysis, and admit the extraction of meta-networks from texts, and attributes of the extracted networks using DIME/PMESII cross classification. A combination of machine learning, rule based, and statistical algorithms will be used.

Research Approach

The application of DNA techniques to a large complex system, such as al Qaeda or the U.S. Army, entails a series of procedures. First, one needs to gather the relational data. One approach for doing this is to extract relations from a corpus of texts including open-source items such as Web pages, news articles, journal papers, stockholder reports, community rosters, and various forms of humint and signals intelligence (sigint). Second, the extracted networks need to be analyzed–that is, given the relational data, can we identify key actors and sub-groups, points of vulnerability, and so on? Third, given a set of vulnerabilities, we want to ask what would happen to the system were the vulnerabilities to be exploited. How might the networks change with and without strategic intervention? The CMU CASOS group has developed an interoperable suite of tools, shown in Figure 1, that acts as a chain to extract networks from texts, analyze these networks, and then engage in what-if reasoning.

This tool suite takes into account multi-mode, multi-link, and multi-time period data including attributes of nodes and edges. This toolset contains the following tools: AutoMap for extracting networks from texts, Organizational Risk Analyzer (ORA) for analyzing the extracted networks, and DyNet for what-if reasoning about the networks. This proposal is to improve the data extraction tool AutoMap and to adjust the analysis tool ORA so it can utilize the new types of data extractable via the improved AutoMap.
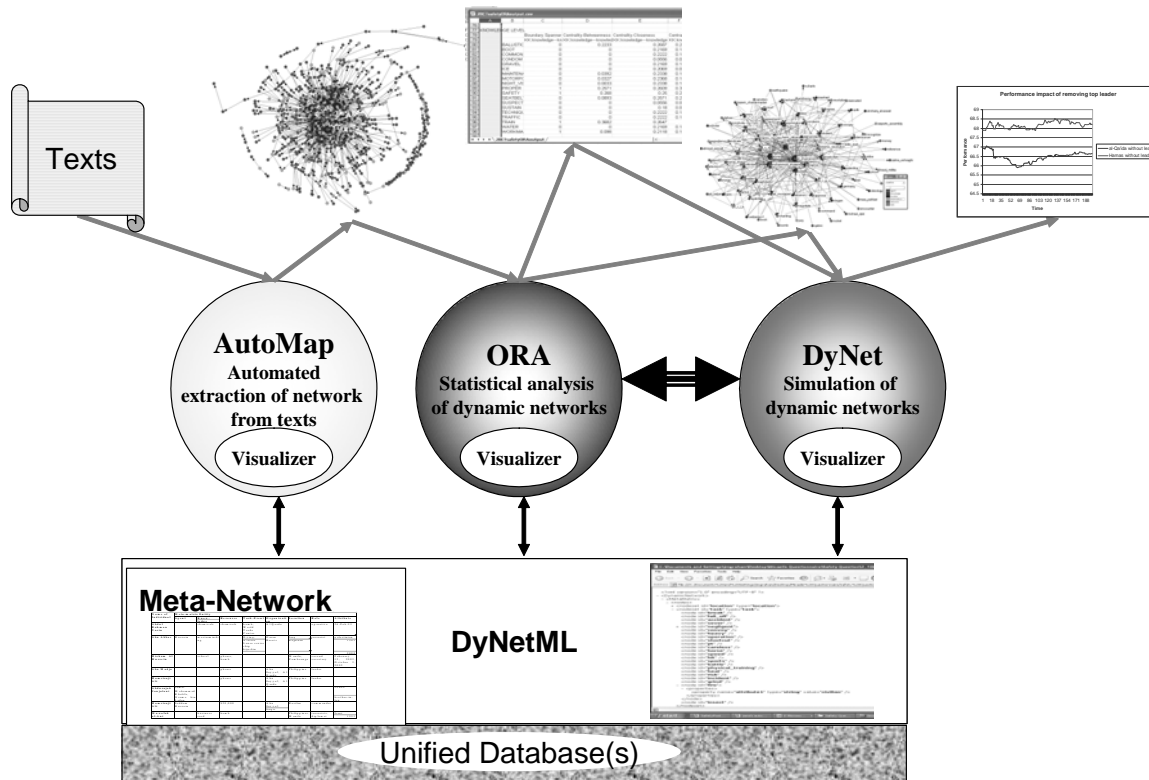
128

Figure 1. *CASOS tool chain.*

Accomplishments

Improved thesaurus construction in AutoMap; in particular, add new techniques for more automated thesaurus construction. We developed a parts of speech processor that characterizes concepts in text by part of speech (POS). Next, we will expand this by creating rules that then map these parts of speech to the ontological categories used by AutoMap. We found that conditional random fields was the best model to use. We also found that failure to clean texts (e.g., by leaving in typographical errors and extraneous punctuation), leads to artificially high classification scores. Most tools doing POS are much worse than claimed because of the overability to find unique cases due to typographical errors and punctuation.

Add a latent semantic analysis (LSA) component to AutoMap to do more automated ontology creation. After much exploration, we realized that LSA was best used in between rounds one and two of AutoMap coding; that is, after the content analysis and before the semantic networks and meta-network extraction. So we developed an LSA tool, added it to ORA, and then built a script to run AutoMap ORA-LSA to reduce the selection of texts, and then rerun AutoMap. More work needs to be done on how this process can be used to help Soldiers rapidly process huge quantities of texts.

Add an LSA component to pre-classify texts for processing by AutoMap. This is finished; it is part of ORA.

Enable AutoMap to meta-tag texts coded so that the meta-tags can be used by other systems. We are working on this, but this is not completed. A key issue is that there are not agreed upon standards for meta-tagging of texts.

Add tools to AutoMap to extract information on time of activities discussed in texts and attributes of actors, events and locations. Time discussed in a text is extracted, but temporal relations are not; however, extracting time turns out not to solve the problem of wanting to tag events and interactions by when they occur. Further exploration is needed for auto-identification of appropriate time chunks and extraction of time from meta-data.

Improve speed and scalability of AutoMap. Explore whether it is feasible to improve the speed at which AutoMap processes texts through multi-threading or parallelization. We are currently doing an analysis to identify points for speedup. The trivial part—run across multiple platforms with one text per platform—is done.

Explore the use of machine learning techniques to improve the text-mining approach used in AutoMap. We have explored hidden Markov models and conditional random fields (CRF), and we will be looking at additional models. It is important to recognize that machine learning (ML) techniques for text require 2D linear processing, which is very distinct from the 3D non-linear techniques used for network data. So totally different forms of hidden Markov models and CRF are needed for AutoMap and ORA. A second key point is that it takes months to develop an ML model and then optimize it for a type of data; then you run it on a test set and "learn" an "application model." Then the application model can be used as a classifier to interpret texts in the future. The application model is what is integrated into AutoMap. We are working on application models for parts of speech classification, classification of concept by position and part of speech in the DNA ontological categories, anaphora resolution, list processing, and identification of negative relations.

Make the ontology used in ORA flexible (it is currently predefined and rigid). This will be done by enabling the user to self-define relative to the who, what, when, where, how, and why categories. This should be finalized by June 1, 2008.

Create merge and tracking routines for thesauri and databases constructed from different sources. We have merged ours and are creating a scheme for tracking and merging others. We will work this coming year on the technology to support this plan.

Explore linking AutoMap and ORA to streaming or regularly changing data streams. We began by trying to link to communications data; ORA can handle it. AutoMap is not needed. Next we plan to link to continual Web scrapers.

Augment DyNetML to handle meta-data information. This has been completed. We also are building a DyNetML checker and relaxed-NG specifications, so others can check their DyNetML as well.

Link AutoMap and ORA directly to the MySQL database. AutoMap can output data to a structured query language (SQL) compliant database and ORA read from it. Next we will define

our database specs and enable AutoMap to read from the database as well. We are doing this: (a) to handle massive amounts of data and (b) so that Soldiers can adjust the thesauri from either AutoMap or ORA and the changes will propagate. (e.g., I changed my mind; I want to combine all trucks, cars, and jeeps as "autos" and I want the change to propagate for both analysis and coding; or, I just found a new alias for Bin Laden and I want to code for it and combine in my data all "actors" with both the original name and the alias to the same actor.

Update help and add simple wizards to ORA and AutoMap to facilitate tool use by the military. All of ORA help was just updated and made more intuitive. We are doing this for AutoMap now. This will be released in mid-June as tech reports and new user guides.

Test tools by analyzing data collected from:
- The Web for information on various terrorist groups,
- Enron corpus for email,
- Information collected at various military experiments, particularly at the battle lab, and if possible that collected using CPOF, and
- Other data corpi that come available to use from the Department of Defense

Note. This process is ongoing. If you have data you want us to process, please let us know. We would be happy to jointly collaborate on a project.

Provide the tools and data from the battle lab regularly to West Point for use by cadets under Lieutenant Colonel John Graham. We have provided all tools to the USMA and they are being used.

Utilize feedback from cadets, intelligence analysts, and others to improve user-interface. This is ongoing.

Contributions to Basic Science

Interest in social network science is continuing its decades-long rise mainly due to its usefulness in an assortment of broader research disciplines, such as organization behavior, sociology, computer science, and forensic science. A key problem is that network science tends to be applied primarily to data collected via surveys or from membership manifests. This means that vast quantities of ethnographic and qualitative data go unanalyzed, as they can currently be processed only through time-consuming and error-prone manual coding. The tools proposed will make it possible to assess and analyze more of the information in these texts, reduce coding barriers, and facilitate rapid automated and semi-automated coding.

Potential Army/Military Applications

Key military applications include: assessment of hot topics under discussion in foreign countries, political and social attitude assessment, identification of key opinion leaders and changes in their opinions, coding of reports by commands and leaving them as reports to future commands in an extensible and usable form, assessment of degree of misinterpretation of commanders' intent, assessment and cross-classification of rules of engagement.

Automatically-Learned Representation of Complex Task Environments to Support Training and Collaboration

Contract #: W91WAW-07-C-0013
Institution: Knowledge Analysis Technologies, LLC

Contract Dates: 04/09/2007 to 04/08/2009
PI: Peter W. Foltz

Problem(s)/Research Question(s) – Due to the inherent complexity of distributed collaborative team environments, effective measurement is beyond the powers of traditional task analysis. We are developing theory, methods, and a generic automated tool that will model, assess, accelerate, and aid the experience-based acquisition of human knowledge and skills needed in distributed collaborative tasks.

Technical Barrier(s) – Conventional approaches to measurement in task environments requires labor intensive efforts in developing task, cognitive, or user models. In complex team environments, these models become quite burdensome to develop as well as quite fragile. This limits the ability to develop effective measures of team performance as well as systems to evaluate, train, or monitor teams.

Significance/Impact for Basic Research – This project addresses two interlocking basic research problems. The first is how to use modeling approaches to improve the efficiency and effectiveness of distributed net-based collaborative problem solving and team performance operations. The second is to improve techniques for characterizing performance within complex problem solving situations.

Potential Transitions – The potential for technology based on the underlying theories and methods developed under this project can lead to software for automated team and individual performance analysis within training systems. These can include leadership training systems at the Army War College (AWC), Fort Leavenworth, automated After Action Reports (AARs) systems for convoy training, such as at Fort Lewis and the National Training Center (NTC), and generally for any command and control and/or complex team simulation training systems.

Overview

We are developing theory, methods, and a generic automated tool that will model, assess, accelerate, and aid the experience-based acquisition of human knowledge and skills needed in distributed collaborative tasks. The automated task analysis tool relies on a new, fundamental extension of the mathematical/computational theory *Latent Semantic Analysis* (LSA, originally invented by the principals of the project) that is already in use in automated Army training, assessment, and performance aids. However, instead of learning from massive bodies of expository text as in traditional LSA, the extension learns from extensive records of the sequential communications and actions of actual collaborating groups.

The technology is especially promising for tasks and situations where sufficient characterization of the context, knowledge, and skills involved is beyond the reach of coding by subject matter experts (SMEs) and is captured by manually constructed models because of the

number and fine granularity of the variables involved and/or the rate at which they change. Examples of such tasks are decision-making and planning by distributed command groups, team control of unmanned vehicles, aircraft control, and trouble-shooting of electronic weapon systems.

The machine-learning approach allows the system to emulate and predict both nearly repetitive and novel human actions in the same domain as that in which the system was trained. It does this by measuring the similarity of new events to old, where events are the combination of potentially infinitely variable situations with potentially infinitely variable human knowledge, decisions, and actions. It is able to capitalize on communication data (both content and patterns of who speaks to whom, when), participant actions (button presses, commands, decisions), and environment and system states.

One of the critical aspects of such a research endeavor is the acquisition of appropriate data. The requirements include that we have data from complex problem solving or planning tasks. This requires that there be multiple team members interacting in a domain in which each person has different amounts of relevant information and the team is working together to accomplish certain goals, while still dealing with different levels of uncertainty and lack of information. Over the course of this 3-year research project, we will be developing and testing our approach on different datasets that meet these requirements. The planned datasets are (1) asynchronous discussion data generated during an Interagency Process Simulation conducted at AWC, (2) communications and control actions made by teams of students at the Naval Postgraduate School (NPG) using a command and control simulator called C3Fire, and (3) team communication and control actions to be collected at Fort Lewis using the Ambush! simulator for convoy operations training. These datasets have been either already obtained or will be obtained through existing collaborations. In the third year of the effort, we plan to collect additional data using one of the above distributed simulations in which we will experimentally vary the conditions and goals as well as test the effects of automated performance and moderation.

Success of the research endeavor is being evaluated by the ability of the system to dynamically predict the probability of outcomes of new trials as they progress, score individual actions according to their predicted effects thereon, assess the overall task-relevant knowledge and collaboration skills of both individual participants and groups as a whole, and to accelerate task success by automatically generated warnings, critiques, and coaching of play. Such measures may be thought of as objective assessments of fundamentally subjective human abilities, and in most cases intuitive rather than logically reasoned because of the complexity of both the automated governing of the game conditions and the human participation, interpersonal relations, communication, and influence factors involved.

In all cases, knowledge on the part of the model will be acquired solely by machine learning from previous games with other humanly constructed information inserted only for the experimental purpose of varying the task. No task-specific SME or task-repository information will be involved, the point being to perfect and evaluate the automated task analysis and assessment methodology without confounding the basic research questions with other variables.

Given success, further research and development would presumably magnify the method's effectiveness in cases where hybrid approaches are feasible.

The idea here is that such a modeling ability, if successful, will be applicable to machine assessment, instruction, and aiding of real-world group tasks such as command, control, and mission planning, where the complex background and context data for the machine learning model comes from automatically captured human (network text or oral) communications and system states in actual usage of simulator trainers or operational performance rather than (or in addition to) expertly constructed assessment and training systems.

A potential additional capability of the system would be to automatically generate AARs. These might take the form of transcripts and quantitative estimates of the value of every communication and action in the processes and its probable effect on outcomes according to the model. Such a report might be presented as a timeline, annotated with a structural equation-like diagram of the influence path of each human communicative event.

Research Approach

Over the course of three years, we plan to obtain the datasets mentioned above, research and develop methods for automated task and communication analysis, and evaluate their predictive performance. In the final year of the project, we will run an additional experiment in which we examine the effects of incorporating the technology to serve as an automatic moderator of teams. We will use iterative testing and design to ensure high usefulness and usability of the automated moderator (see Landauer, 1995). The research plan is as follows.

Year One

The initial focus of year one was obtaining the datasets and initial development of the modeling technology. The key tasks for year one included:

1. Obtain datasets. We will obtain data from AWC, C3Fire, and Ambush! events. We will further obtain performance measures associated with the team data. These measures will be a combination of objective performance (mission success, kills, time, optimal path deviations, task time) and subjective SME and participant ratings (performance ratings, situation awareness, communication quality, critical errors). These ratings will examine both teams and individuals, as well as look at team process and the content variables.

2. Develop technology for Latent Problem Solving Analysis (LPSA) of team interaction streams. We have adapted the existing LSA text-based technology to LPSA for team data. Using available data, we plan to generate the high-dimensional LPSA problem spaces.

3. Further refine research methods for analysis of content of spoken and typed communication. We will determine types of language indicative of good and poor performance (e.g., loss of situation awareness, planning, use of acknowledgements to indicate information received). Once these discrete types of language have been

automatically identified by LSA, they can be represented as additional discrete elements within LPSA. This will permit the automatic association of team actions with communication events (to be performed in year two). We will test the performance of the LSA communication analysis methods by correlating the LSA predictions of language types to the objective and subjective team performance measures.

Year Two

1. Evaluate LPSA representations. We will use a training and test methodology in which we develop a LPSA model based on a subset of the data. We will then use the remaining data as test data in order to predict performance scores using the derived LPSA model. We will iteratively refine the LPSA methods as well as test the boundaries of how much training data is necessary for effective and generalizable LPSA representations. Evaluation of performance of the LPSA models will be made against the objective and subjective team and individual performance ratings. We will evaluate the effectiveness of LPSA in measuring team process as well as information content, and will further evaluate the amount of task relevant data required for accurate performance. This will give an idea of the level of granularity in which we can determine when a team is having difficulty. For example, does LPSA require at least 1 minute, 3 minutes, or 5 minutes of prior events in order to sufficiently characterize the quality of performance? In this analysis, we can determine LPSA's performance by providing different size windows of data and testing its ability to predict the next event or action.

2. Integrate LSA-based communication predictions within the LPSA model. Because LSA is able to code the actual content of individual team member utterances (e.g., planning, acknowledging, uncertainty, etc.), these will be integrated as discrete events within the LPSA representation. We will develop and refine the methods for predicting important communication events and integrate them into the LPSA technology. We will evaluate the types of communication events that are most effective at improving overall LPSA prediction performance.

Year Three

1. Develop an integrated LPSA-based moderation tool. The LPSA tool will incorporate data from the communication, action, and event streams. Assuming that LPSA will provide sufficiently accurate predictions of team performance during complex problem solving situations, we will develop methods and an interface that will alert team members (or potentially, commanders) when a team's performance is failing. Failures in performance will be determined based on LPSA's automatic derivation of the team models. We will develop focused feedback that should return errant teams to the appropriate contexts before major errors should occur.

2. Evaluate the LPSA-based moderation tool. In collaboration with one of the three data providers (most likely either AWC or the NPG), we will design a controlled experiment testing the moderation features. Different teams will receive a Knowledge Post-based interface that will either provide LPSA-based moderation or no moderation. We also will

vary other conditions, such as the level of knowledge and skills of the teams, to test the effects of moderated performance on varying abilities.

3. Tests of generalizability across datasets. Based on the analyses performed on the three diverse datasets, we will determine the most appropriate domains for the application of LPSA modeling.

Accomplishments

The research has been under way since April 2007, thus, at the time of this report, we are still in year one.

Data collection. We already have data from AWC and a small amount of data from C3Fire. We have started discussions with additional researchers of C3Fire (Jeffery Thomas, Army Research Laboratory, ARL) who have larger C3Fire datasets. By leveraging off of data collection done as part of the Defense Advanced Research Projects Agency, Competence Assessment and Alarms for Teams (DARPA-sponsored DARCAAT) program, we have obtained 100 hours of communication data from Ambush! but have thus far only received about 5 hours of event data. We are currently negotiating to obtain additional event data.

Develop technology for LPSA analysis of team interaction streams. We have adapted the existing LSA text-based technology to LPSA for team data. This involved changing our existing code-base to process non-textual data.

Further refine research methods for analysis of content of spoken and typed communication. Using data and methods developed under the DARCAAT program, we have continued to refine our communication analysis methods. This research has included examining team communication analysis methods applied to Ambush! and NTC convoy team data and SME ratings of situation awareness, battle drills/action drills, standard operating procedure, command and control, and overall team performance. Initial results show strong ability to predict these SME metrics based on communication.

Contributions to Basic Science

The project is directed at two interlocking problems. The first is to improve the efficiency and effectiveness of distributed net-based collaborative problem solving and team performance operations. The second is to improve techniques for characterizing performance within complex problem solving situations. These problems have traditionally been approached by task analysis methods based on interviews with SMEs, observations of expert performance, and manually constructed specifications and organization of tasks and their associated required knowledge bases. The innovative approach responds to the burdensome time, expense, and intrinsic limitations of conventional techniques when the environmental context is so complex—composed of large or an infinite combination of influences—as to profit from the application of automatic computational techniques. There are two critical differences in the approach relative to other approaches to modeling complex tasks. They are: (1) We capitalize on large amounts of team data, including analysis of the content of what is being expressed, the patterns of

communications, and actions of participants, and (2) The techniques for characterizing performance permit automated derivations of task analyses without tedious hand-coding of tasks.

Potential Army/Military Applications

Monitoring, assessing, and moderating teams of decision-makers in complex military environments requires effective tracking of individual and team performance. The research advances our understanding of modeling performance in complex task environments and permits analyses of complex communication and action patterns. It can therefore be applied in any area where automated task analysis and performance measurement is needed in interactive team situations.

The potential for technology based on the underlying theories and methods developed under this project can lead to software for automated team and individual performance analysis within training systems. The toolset can show integrated concepts and approaches to analyzing communication and control actions, providing predictions of performance, and automated feedback. Such a system could provide automated approaches to:

- Detecting individual and team errors and knowledge gaps,
- Predicting cognitive workload and situation awareness,
- Predicting overall team performance,
- Predicting failures in team process and inappropriate contexts for the mission,
- Monitoring group dynamics,
- Providing automated AARs, and
- Assessing and training leadership.

The development of products based on this research could provide complete automation of team and individual performance measurement. Products can be developed that may be used to facilitate analysis in a host of applied settings, including the assessment of teams in combat information centers, air-defense systems, and remote battlefield command, control, communications, and intelligence (C^3I) centers. In addition, because LSA works for any language, the techniques and tools can be used for monitoring and assessing team communication and actions of opposing forces speaking other languages (as well as their actions).

References

Landauer, T. K. (1995). *The trouble with computers: Usefulness, usability and productivity.* Cambridge, MA: MIT Press.

The Impact of Prior Knowledge on Trust Development in Geographically Distributed Teams

Contact #: W91WAW-07-C-0050   **Contract Dates:** 07/19/2007 to 07/18/2008
Institution: The Boeing Company   **PI:** Mark Handel

Problem(s)/Research Question(s) – Does prior knowledge (of team members, team processes, and individual roles) help a geographically distributed team come together more efficiently?

Technical Barrier(s) – The primary barrier with this research is access to appropriate teams. Most of the work done previously was done in an academic setting, using artificial tasks or student project teams. These do not have the same set of constraints and established processes within which real-world teams generally work.

Significance/Impact for Basic Research – Training and team selection may be a critical factor in producing high performance teams.

Potential Transitions – This research can be applied directly to improve training and team membership selection processes.

Overview

We know a good deal about how individual factors such as extroversion and cultural factors such as obedience to authority impact team formation. Unfortunately, by the time individuals start working in teams, these factors are more or less constants. But, what of factors that can be trained prior to team formation or used to determine team make-up? In particular, we look at four factors: prior knowledge of team members; process knowledge; role familiarity; and leadership style. How does each of these factors impact the development of trust within a team? If these do have an impact on the team, it makes sense to target appropriate training and adjust team selection factors.

Research Approach

The basic approach to answering the research question is taking a set of real teams from the workplace and surveying them at several time points during the formation and work of the team. In addition to the surveys, we also are doing in-depth interviews with selected members of the teams, to make sure there are not critical factors that are being overlooked.

The research looks at four factors that influence team trust formation: prior work experience and reputation, role familiarity, process knowledge, and team leadership style. The first three factors are attributes of all members of the team, while leadership style is an attribute of the team leader. The leadership style is determined through survey questions for all members of the team. Role familiarity is essentially how well the team members understand why they are on the team, what they bring to the table, and what the other team members contribute. Process knowledge is the understanding of process (e.g., field manuals, business process documentation), and the tools used to execute the process (e.g., any specialized computer systems, spreadsheets,

forms). This work primarily focuses on teams that are executing established processes. The details may change, but the big picture remains more or less constant. Within the research site, we have identified proposal capture teams as a good example of this.

The proposal capture teams have many good attributes for this research. First, they are usually geographically distributed. Secondly, there is an established business process for analyzing requests for proposals (RFPs), and responding with a completed proposal. Some aspects of this process are well documented; for instance, developing a cost proposal as opposed to determining an overall technical direction. This provides a framework in which most people are expected to understand their role, but still provides enough ambiguity and variation to get interesting results. Thirdly, the teams form on an ad-hoc basis in response to individual RFPs. The result is that some people have worked together on previous teams, and others are new to either the team and/or the process.

By surveying the team as it is forming, a baseline is established. This allows us to see what the baseline conditions for all of these factors are, and how they change over the course of the research period. The second survey is to see if any of the factors have changed, reflecting greater understanding of the team and the members' roles in the teams. The second survey, several weeks after the team has been working together, is also when outcome measures are taken, looking at individuals' trust in the team, their perceptions of team cohesiveness, and the leader's perception of the team's effectiveness. In between the first and second surveys, there are in-depth interviews with selected participants. This will help to give examples and stories that explain the underlying roles of the factors identified.

Both qualitative and quantitative analyses will be performed on the data gathered from the surveys and interviews. The initial unit of analysis will be the team. We will use structural equation models to test which scales most strongly influence initial trust, final trust, and team performance. In addition, we will use the data about prior knowledge of individual teammates to build a social network model. We will use these social network models to compute measures of connectedness that will be used as additional factors in the structural equation models. We will use the data from the qualitative interviews to understand better the nature of leadership in the development of team trust. In addition, the qualitative interviews will help us understand the nature of technology and communication technologies in particular. Although not the primary focus of this work, this may have impacts on trust that need to be explored.

Accomplishments to Date

This work is on-going; however, in the first part of the project, the basic surveys were created. This involved creating a new instrument to measure process and tool familiarity. This instrument was validated against a number of people within the Boeing Company and refined based on their feedback and statistical analysis of the results. In addition, we have been actively soliciting teams from within our business unit for participation in this effort.

Contributions to Basic Science

Much of the research around trust formation in teams uses a model that posits there are two types of trust at work in most organizational teams: an emotional trust and a cognitive trust (model from McAllister, 1995). This work extends the notion of cognitive trust from previous works (Rocco, Finholt, Hofer, & Herbsleb, 2001), in which cognitive trust was defined to be solely a measure of competence and reliability. If successful, we will show evidence for the existence of another aspect to cognitive trust—organizational understanding. Organizational understanding is suggested in part from prior research on transactive memory (Hollingshead, 2001).

Potential Military Applications

Organizational understanding is not necessarily a trait of an individual, but rather qualities that concerted management efforts can easily put into place, change, and improve. If this organizational understanding construct proves to be an important factor in trust development, there is a wide body of work on how to train for these factors. This provides an immediate opportunity for military application in terms of changing training plans. A slightly more complex application would be to use the evidence about how team composition (in terms of prior experience with team members) can be used to create better teams: for instance, how many new people should be on a team, versus how many experienced people?

Future Plans

As explained in the accomplishments, the baseline surveys have been created and validated against; we are in the process of identifying appropriate teams and administering the surveys. After the data collection, the analysis will be done. Because of the design of the experiment, it is hard to say much during the data collection phase of the project.

References

Hollingshead, A. B. (2001) Cognitive interdependence and convergent expectations in transactive memory. *Journal of Applied Psychology*, *81*(6), 1080-1089.

McAllister, D. J., (1995) Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations," *Academy of Management Journal*, *38*, 24-59.

Rocco, E., Finholt, T. A., Hofer, E. C., and Herbsleb, J. D. (2001, April). Out of sight, short of trust. Presentation at the Founding Conference of the European Academy of Management. Barcelona, Spain.

Detecting Changes in Dynamic Social Networks

Army Project #: 611102B74F   **Contract Dates:** 10/01/2006 to 09/30/2008
Institution: United States Military Academy   **PI:** Major Ian A. McCulloh

Problem(s)/Research Question(s) – Social network analysis is a powerful tool for understanding and modeling human behavior, because it models the complex and highly dependent relationships between people. This project investigates the feasibility of using email data within a military organization to study social networks and provide new management tools for the information age. In addition, new methods are developed to statistically detect changes in social networks in real time to provide early warning and inform organizational policy.

Technical Barrier(s) – In 2005, the National Research Council report on Network Science, recommendation number 1 states, "*The federal government should initiate a focused program of research and development to close the gap between currently available knowledge about networks and the knowledge required to characterize and sustain the complex global networks on which the well-being of the United States has come to depend.*" Traditional statistical methods usually require necessary assumptions that are not satisfied in relational network data. In order to explore valid statistical methods for networks, suitable data must be used to validate methods. There is a significant lack of social network data that has continuous time stamps for this type of research. When similar data has been collected, it usually lacks important information about organizational activity to validate change detection.

Significance/Impact for Basic Research –

1. Predictive Models in SNA - The National Research Council (NRC) in a recent report on Network Science identified a lack of understanding in the stochastic behavior of networks. They further stated that there existed a great need for this understanding in order to develop effective predictive models. This thesis will provide insight into the stochastic behavior of social networks. In addition, algorithms will be proposed that detect subtle changes in a social network. Imagine detecting a change in the social network of a terrorist organization as they increase their communication before they are able to execute their planned terrorist attack. Or consider a civilian company, whose managers can identify major leadership challenges before they affect the productivity of the company. These important tools may become a reality with this research. Furthermore, they will be implemented in commercial SNA software packages.

2. Statistical Algorithms - Statistical process monitoring (SPM) is an area of quality engineering and statistics designed to take sequential samples from a system and detect changes quickly. These same techniques that save industrial manufacturing processes billions of dollars a year can also allow sociologists and psychologists to more effectively identify changes in a social network compared to current methods. This research will not only present easier and more accurate methods of DNA, but they will also save time and money in live-experiments. Through the use of SPM, changes in social networks over time may be detected more rapidly, while reducing false alarms.

3. New Line of Research - Broader impacts of the project include establishing a method for detecting changes across all standard measures used in SNA. This will create opportunities for implementing other more robust algorithms to detect DNA changes even more rapidly. This tool will enable intelligence analysts to more effectively paint the operational human terrain picture. In addition, the Army will be able to use these methods to gain greater insight into operational exercises such as BCTP.

4. Impact On DoD Capabilities - There is a DoD need for understanding the social networks of insurgencies, terrorism, friendly command and control, arms trade, WMD proliferation, and many other applications. The methods investigated under this proposal are designed to model these dynamic social networks and extract important information about when these networks change over time. Understanding and modeling these complex networks is essential to understanding threats and current capability, as well as to predict the future behavior of these networks.

Potential Transitions –

- Improved tools for command and control in a networked environment. SNA has been applied to improve organizational behavior in civilian applications. This research investigates the ability for commanders to apply SNA techniques to command and control.

- Networked communication. We are still investigating the impact of one regiment in the cadet military chain of command having blackberries. Does this improve the responsiveness of the operations staff?

- Intelligence agencies can apply these techniques to monitoring terrorist organizations.

Overview

A relatively new area of SNA research looks at how the network changes over time. This is known as Dynamic Network Analysis (DNA). An early DNA study of a Battle Command Training Program (BCTP) exercise conducted at Fort Leavenworth, KS, revealed valuable information about how battalion and brigade staff members interacted with one another over the course of the exercise. This analysis allowed investigators to improve upon how staff officers were task organized, thereby improving the effectiveness of new transformational Army units. Being able to detect when the network changes is extremely important to this type of analysis. This research will apply methods of detecting changes in commercial manufacturing processes to detect changes in social networks.

In addition to demonstrating the proposed methodology on a variety of existing data sets, this project will also involve the collection of a rich and innovative network data set. Social network data will be collected by monitoring the e-mail traffic of 24 Army officers in a one year graduate program at Columbia University. This data is unique in that when we detect changes in the network over time, we can survey the officers and determine why these changes occurred. We can explore in detail, what causes people in an organization to interact. Is it serving in prior

units, going to the same church, having kids the same age, sharing common interests? Which of these has a greater impact on the interaction between individuals? In addition, this data will be available to others conducting research in the field of SNA and will make a significant contribution to the discipline. With this data, new algorithms will be developed using maximum likelihood estimation to determine when the networks change. Different statistics will be studied on the data set that allow faster detection and identification of changes in the network while minimizing the probability of false alarms. These same algorithms will be automated in a software package and be applied to various SNA measures to rapidly detect and identify when changes occur in a social network. These methods may provide actionable intelligence when modeling insurgent networks. They can also provide a valuable means of identifying potential improvements in friendly command and control networks. The potential applications of this research are endless.

Research Approach

Research funds are used to equip officers in the Eisenhower Leadership Development Program (ELDP) with laptop computers in exchange for their agreement to allow the principal investigator to monitor their sent email traffic. In year two, the ELDP members were given blackberries, since the Army now resources them with laptop computers. Social Network email data is collected over the course of their year-long program. The members in the ELDP were also surveyed to collect social network data on the group. The email and survey data collection methods were compared. Different software approaches to collecting real time email data were developed and compared. Change detection methods traditionally applied in quality engineering and manufacturing was applied to the ELDP social network to detect change in real time. Technical barriers were addressed with innovative software, and the unique data set. When changes were detected in real-time, ELDP members were interviewed to determine the causes for the detected changes.

Accomplishments

- E-mail is an excellent, unobtrusive source of social network data. This data can provide commanders great insight into the organizational behavior of their command.
- E-mail data collected directly from the DOIM server is more reliable, complete, and efficient for data than client side collection methods.
- Statistical process control charts used in manufacturing have been shown to effectively detect significant changes in the ELDP program as well as in an Al-Qaeda data set.
- ARCENT has used methods developed in this project to improve the organizational efficiency of the command.

Contributions to Basic Science

This research has significantly expanded the body of knowledge about complex social networks. This project has developed a unique and valuable data set necessary for a deeper understanding of network dynamics. In addition, we have made significant progress toward understanding the probability mechanics of socially networked human communication. This is precisely the basic scientific focal point of the NRC 2005 report on network science. Future

research in this area is aimed at better understanding the ergodic, highly dependent intricacies of relational network data.

The change detection methods developed in this project have been implemented in the Organizational Risk Analyzer (ORA), which is a proprietary software developed by the Center for Computational Analysis of Social and Organizational Systems (CASOS) at Carnegie Mellon University. This software is available to all DoD agencies under a free license. It is widely used by many combatant commands for intelligence. This research has also provided academic research opportunities to over 30 cadets and 6 faculty at the U.S. Military Academy.

Potential Army/Military Applications

- Commanders can use the email data collection methods to study the organizational behavior of their command.
- Commanders can detect significant organizational change in their command and receive early warning of potential command problems and issues.
- Intelligence agencies can monitor terrorist, and drug organizations for significant change in behavior in real time.
- Policy decisions can be evaluated by looking at the historical change points of organizations and evaluating the direction and magnitude of their changes.

Future Plans

We intend to expand this research to involve more socially connected subjects. The subjects will still include ELDP members, but will also involve cadets and faculty. With the larger pool of human subjects, we will be able to evaluate how well findings generalize to other groups. In addition, we will collect additional evidence of the success of Social Network Change Detection (SNCD). The probability mechanics of social networks will be explored by modeling the studied population with exponential random graph models and network probability matrix models. The population will also validate models constructed in social simulation software. These validate models will allow investigators to explore a greater range of possible social changes and evaluate multiple methods of anomaly detection.

Successful Collaboration at a Distance

Contract #: W91WAW-07-C-0060   **Contract Dates:** 08/13/2007 to 08/12/2010
Institution: University of Michigan   **PI:** Gary M. Olson

Problem(s)/Research Question(s) – While modern communication and computing technology make it possible for groups or teams to work together when their members are in different locations, it is by now well known that it is difficult to do so. A variety of factors have been identified as possible reasons for this, but research is needed to clarify the role of these factors, including the relative importance of each.

Technical Barrier(s) – Identifying why it is difficult to work at a distance has required the collection of large amounts of data in both the field and the laboratory. We have developed a set of working hypotheses about what factors affect the success or failure of long-distance collaborations. While the total set of factors is quite extensive, they fall into five broad categories: Collaboration Readiness, Technology Readiness, Common Ground, Management, and Nature of the Work. Specifics about the full set of factors can be found in Olson, et al. (2008).

Significance/Impact for Basic Research – Many researchers are interested in this problem, but to date, no comprehensive account exists.

Potential Transitions – We hope to provide those with goals of carrying out successful collaborations at a distance with specific guidelines for how to maximize the chances of success.

Overview

This research seeks to understand the factors that contribute to success or failure in geographically distributed projects. In earlier work, we generated a series of working hypotheses about factors that are important in geographically distributed projects. In this project, we seek to verify these working hypotheses and determine their relative importance.

Research Approach

We are carrying out four lines of research in this project:

1. Research efforts in the field of teams working at a distance.
2. Laboratory experiments of selected factors,
3. Computational models of the task involved in the laboratory experiments, to help select which variables to study empirically, and
4. The construction of a collaboratory online wizard that will be a resource for those about to embark on a geographically distributed collaboration, to assess the risks involved in their project and receive recommendations as to what to do about those risks.

Accomplishments

This project has just begun, so we have made only preliminary progress. Our progress to date is as follows (and corresponds to the number items under Research Approach above).

1. We have been carrying out interviews with teams in companies that are attempting to work together at a distance. We have gathered 120 interviews from 16 sites so far. These are in the process of being transcribed so we can embark on further analyses.

2. We have designed a series of laboratory experiments of geographically dispersed team members. During the summer of 2007, we conducted pilot research on the task and the experimental design, and in the fall we embarked on the collection of systematic data.

3. The computational modeling is a future project.

4. We have made considerable progress on the computational wizard and have populated it with content drawn from our earlier work. We are beginning user testing of the interface and the survey contents. We expect to be able to deploy this for actual use by the late winter or early spring, 2008.

Contributions to Basic Science

Given how much interest there is in questions of the social and technical barriers to successful collaboration at a distance, the work in progress could have far ranging implications for a number of science areas.

Potential Army/Military Applications

Our findings should be of great value to the Army and the military since geographically distributed teamwork has emerged as a major thrust across all sectors.

Future Plans

Each of the four areas described above will proceed as planned. The only area not already underway (computational modeling) will be engaged shortly; otherwise, we will continue with our field interviews, will carry out one or more laboratory experiments, and will test and then deploy the collaboratory wizard.

<div align="center">References</div>

Olson, J. S., Hofer, E. C., Bos, N., Zimmerman, A. Olson, G. M., Cooney, D., et al. (2008). A theory of remote scientific collaboration (TORSC). In G. M. Olson, A. Zimmerman, & N. Bos (Eds.) Scientific *collaboration on the Internet.* Cambridge, MA: MIT Press.